

nature

20 April 1978

NATO looks at its science programme

THE Science Committee of the North Atlantic Treaty Organisation (NATO) is 20 years old and has been holding a large conference in Brussels to celebrate the birthday. Official delegations from the 15 member nations and various other scientists, administrators and hangers-on have been regaled with two and a half days of a curious blend of science, opinion and diplomatic niceties. Science in the research sense was disposed of in half a morning during which rather more than 300 pages of preparatory papers had to be digested and discussed; the remainder of the time was devoted to more general problems of science and technology and its relationship to society, international relations, education and demography, decision-making and so on. The technological centrepiece was a trans-Atlantic computer link-up which allowed the audience to see the broodings of a computer at Case Western Reserve University on world futures up to the year 2025 in four- and five-significant-figure glory with never an error bar in sight.

What is NATO doing in science? For nearly ten years of the alliance there was no science, at least at the civilian level. Then in 1956, after a backward glance at Article II of the North Atlantic Treaty, which spoke about "strengthening free institutions . . . and promoting conditions of stability and wellbeing", the so-called "Three Wise Men" started to prepare the ground for scientific collaboration. There was at the time a certain unease at the state of the alliance, and also at the state of science (which the launching of the first sputnik in 1957 helped to bring into clear focus). Heads of government of the alliance agreed shortly thereafter that "the full development of our science and technology is essential to the culture, economy and political and military strength of the Atlantic community . . . we must build the established tradition of the universality of true science". In March 1958 the science committee first met. For 20 years it has been occupying itself with, amongst other things, supporting international fellowships, collaborative research grants and advanced study institutes. Last year it spent a little under \$10 million; in recent years its expenditure has been declining in real terms.

Now the things that people say about NATO's involvement in science are that an organisation that is ostensibly military has no business getting into civilian science, indeed will use this support as window-dressing; that the "universality of true science" is no respecter of frontiers and is as likely to require collaboration with the Soviet Union, Sweden, Japan or Australia as with another NATO country; and that not only are there many developed countries not in NATO, there are also a very large number of developing countries

who are excluded from the benefits. In short, that a NATO science programme is a somewhat unnatural concept.

With which it would be very easy to agree if there were a large bureaucracy in Brussels, obsessed with protocol and the correct ordering of things. But somehow the science programme has managed really to foster interchanges, raise the standards of advanced education and simply enable more people to meet more people on a budget smaller than some individual research projects in member states. Whatever one's reservations may be about the geographical confines within which the programme has to work (the Advanced Study Institutes are not quite so confined) the fact remains that within these limits, the programme is a model for how international collaborative science can work. Partly the restricted scale helps, there is no doubt. And partly the very fact that distinguished scientists make up the committee is of great value.

There are, even so, questions that can rightly be asked after 20 years. One of them is the place of industry in the scheme of things. There were very few industrialists at the birthday celebrations—understandably so in view of the usual industrial concern that its own research be protected against prying eyes. But the sort of things that were talked about, and that will probably continue to be talked about in NATO circles, have a growing need of the industrial point of view. Another question is how defence science fits in. Up to the present NATO has scrupulously kept its civil science free from defence applications (except in so far as almost any subject of study can have its defence side). Although in many ways excellent, this policy adds to the isolation of the defence scientist whose opportunities to meet other scientists and talk about problems of common interest are in any case very restricted. A most important job for the NATO science committee could be to poke around the fringes of the security screen and see where there might be areas in which defence scientists could profitably be brought into more contact with their civilian counterparts.

Should "science" now be extended to include the social sciences? Several celebrants wondered aloud about such an increase in scope. The answer should surely be no. If there is good cause for NATO to foster the social sciences on an international basis, it were best that a new office with a separate budget be set up for the purpose. There will otherwise only be suspicion that the science programmes are being drained of resources to help establish the social sciences, and there is enough coolness across the natural science/social science barrier as it is, without inviting more in this way. □



Sorry, for copyright reasons some images on this page may not be available online

Janet Welsh Brown of the Office of Opportunities in Science, American Association for the Advancement of Science, discusses a recent study which suggests that opportunities for women scientists are not yet equal to those of their male colleagues.

WOMEN students and women scientists still receive the persistent message from much of society and from many of their professors that they do not belong in science. In an intensive study of 120 men and women who received their PhDs in the past six years, the majority said that women scientists are treated differently from men—that their work is not taken seriously, their achievements down-graded, their commitment to science always in doubt. The cumulative message may be subtle but it is destructive to women's achievement and self-confidence and it is reflected in lower status, rewards and opportunities.

The study was conducted by the American Association for the Advancement of Science (AAAS) to investigate the participation of women in scientific research. The National Science Foundation supports the AAAS's work which focuses attention on the status of women in science and advises the Foundation on policies and programmes it should adopt.

The goals of the broader AAAS programmes on 'opportunities in science', are to achieve equal access to science education and science careers for all able people. Women, persons of minority status and those with physical handicaps are groups that are seriously under-represented in science though together they constitute the clear majority of the American population.

The 120 women and men in the study detailed the ways in which women are denied models and mentors, excluded from informal professional networks and are sometimes the subjects of improper sexual advances by men in positions of authority over their professional futures. Like women in other fields, they are being hit with a 'backlash' without ever having experienced the 'frontlash' and their accomplishments are often denigrated by the assumption that they only got their positions through 'reverse discrimination'.

At the AAAS we seek equal opportunity not only as a matter of civil rights and human and economic justice but also out of concern that the nation fully utilise its scientific and technical talent. As Senator Kennedy puts it:

The virtual exclusion of women from careers in science and engineering is contrary to our national commitment to equal employment opportunity and weakens our nation's scientific research effort, which needs all the talent the nation can master. To bring about the needed changes, we will need the full co-operation of the scientific community, our academic institutions, public and private employers, and the federal government.

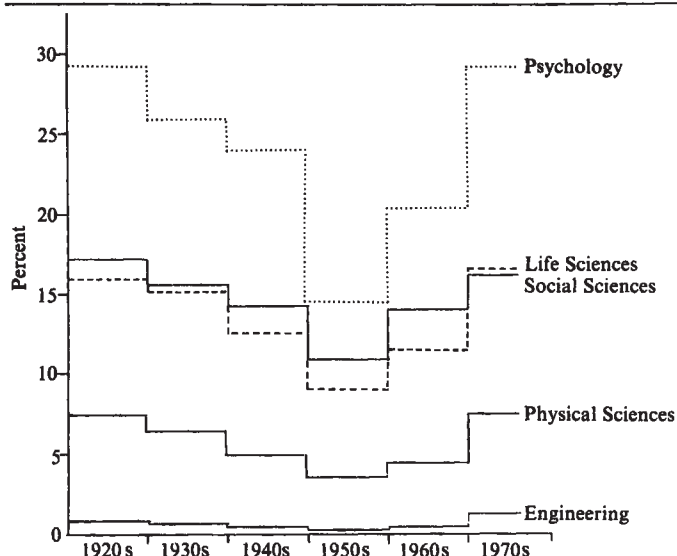


Fig. 1 Percentage of doctorates in science and engineering awarded to women by decade (1920s-1970s). Source: NAS.

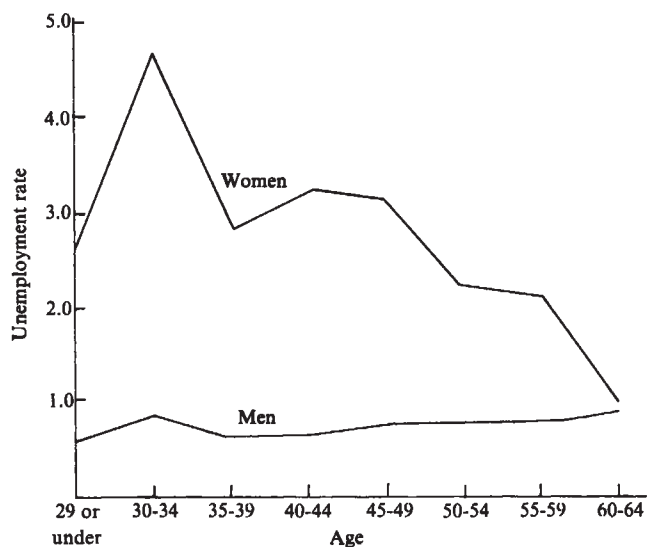


Fig. 2 Percentage of unemployed doctoral scientists and engineers by age (1975). Source: NSF, NAS/NCR, USDL.

To this end the AAAS programme seeks to increase the proportion of women in the sciences, to improve the visibility and status of those women already working in the sciences, and to move more of them into positions of influence and policy-making*. We try to do this within our own organisation of 125,000 members, among our 284 affiliated scientific societies, and in the world where scientists work. Within the AAAS we have sought to have women serving on all the governing bodies and functioning committees, and on the Association staff, and to increase their visibility through participation in the AAAS Annual Meeting. We share these activities with affiliated societies,** and with them attempt to make an impact on hiring and promotion in the universities and on the policies of the federal agencies that fund scientific research and education.

Occasionally, we provide information on request to Congressional staff and sometimes testify at Congressional hearings, but for the most part we work in a collegial way with the scientists (mostly male) who run the federal science establishment. There is wide variation in their interest and commitment to equal opportunity, but most of them are accessible to the public. We use a combination of information, technical assistance and pressure depending on the particular situation and our own resources. Washington is a very communicative town, and a well informed network of feminist professionals widely distributed through the establishment is perhaps the women's greatest resource. Information travels fast.

Our concern is for women in all the natural and social sciences, but we are concerned most of all about the fields in which the sex-ratios are most lop-sided, that is, in the physical sciences and engineering. Figure 1 shows graphically what those fields are. In so far as PhD production is a barometer, this graph also illustrates exactly what is happening for women in science: marked improvement since the 1950s has in the mid-1970s matched the women's record of the 1920s! In most fields, the enrollment patterns of women tend to parallel those of the men regardless of efforts to increase female enrollment in the sciences at the university level. There are still prestigious university departments, like the University of California at Berkeley, where in a faculty of fifty chemists there is not a single woman.

*The goals for minority and physically handicapped persons are the same, though some of the specifics of the programmes may differ.

**These include the two major national organisations for women scientists: the fifty-year-old Graduate Women in Science (Sigma Delta Epsilon) and the seven-year-old, rapidly growing Association for Women in Science.

Though more women are being hired at the lower faculty levels than was true a few years ago in the post-secondary institutions, the salary gap between men and women has actually widened in the past two years. Heavily tenured faculties, declining enrollments and severe financial pressures restrict university opportunities for the young, both male and female, in science as in other areas, just at a time when the proportion of women scientists is growing. They live with the harrowing reality that they are in what is likened to a revolving door which will put many of them out in the street before their careers have barely begun.

There are some signs of improvement, however, and if the universities were not under such financial pressures, some of the improvement might even show up statistically. The proportions of women receiving PhDs in the sciences can be expected to continue its slow but steady increase. Industry, which has employed few women scientists in the past, is now welcoming them, at least at entry level. The private sector has a long way to go but is taking affirmative action seriously, especially since the example of some very expensive law suits. There is genuine opportunity in this area if women scientists will take advantage of it. The women in our study sample who work in industry were well pleased with their positions.

Affirmative action has certainly opened up the recruitment process to some degree. In agencies like the National Institutes of Health which keep track of research applicants and awards by sex, the success rate of women applicants for research grants equals the men's, though the women do not apply in proportion to their numbers—probably because they are not employed in research institutions in proportion to their numbers. Admission to graduate schools is probably sex-fair at most institutions these days, and the most blatant kind of discriminatory statements and behaviour ("I wouldn't ever have a woman in my lab!") is on the decline.

Although very real problems remain and should not be underestimated, I am on the whole optimistic. The most encouraging sign is that the scientific community is facing up to the inequities within it. There is good will and determination in many universities and professional societies and in some federal agencies. Senator Edward Kennedy, who is a power to be reckoned with in national science matters, recently introduced a bill on women in science, meant to make a broad attack on the problems. Such signs of national awareness and leadership augur well, if women and their allies can only maintain the pressure long enough to change those deeply ingrained unscientific attitudes and behaviours. □

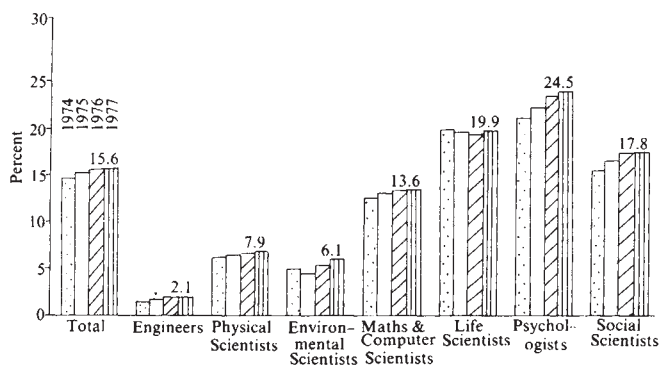


Fig. 3 (above) Proportion of women among scientists and engineers employed in colleges and universities by field (1974-77). Source: NSF.

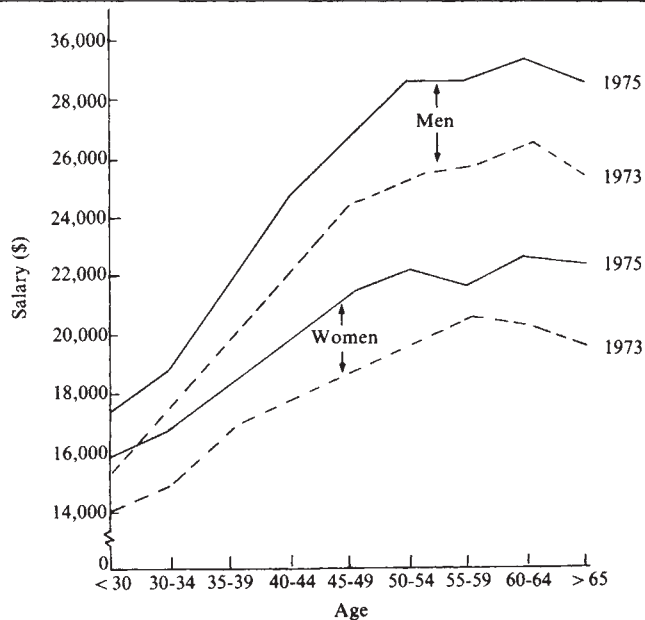


Fig. 4 (right) Median annual salaries of doctoral scientists and engineers by sex and age 1973 and 1975. Source: NAS. (All figures compiled by Betty Vetter, Scientific Manpower Commission, Washington DC).

US Congressional committee rejects Carter compromise on fast breeder programme

PRESIDENT Carter may again be required to use his power of veto in his attempts to terminate the controversial Liquid Metal Fast Breeder Reactor Project at Clinch River near Oak Ridge in Tennessee.

This follows last week's rejection by a key Congress committee of a compromise formula offered by the Department of Energy that would have substituted a long-term design project of a larger breeder reactor, including the study and evaluation of alternative fuel cycles.

The president has, since his campaigning days in 1976, consistently opposed the Clinch River project. However it has been equally strongly supported both by Congress and by the nuclear industry, which argues that the construction and licensing of the reactor is the necessary and logical next step in the development of a commercial fast breeder system.

Last autumn President Carter vetoed an authorisation bill which had been approved by Congress and included money to continue the Clinch River project, although he later signed a supplementary appropriations bill allocating \$80 million for the project, largely because the bill gave him his way on another presidential target, the B-1 bomber.

In his proposed research and development budget for the Department of Energy for 1979, which was presented to Congress in January of this year, the President recommended that the Clinch River budget be reduced to \$13.4 million and that this money be used to terminate the project.

The fossil and nuclear energy subcommittee of the House Committee on Science and Technology, however, which contains a number of nuclear supporters, revised the president's request and inserted provision for \$173 million to be spent on the project, not necessarily for its termination.

In an effort to forestall what seemed an inevitable extension of the conflict between Congress and the adminis-

tration, Dr James R. Schlesinger, Secretary for Energy, proposed a compromise formula last month in which, in return for Congress' acceptance of the termination of the project, the administration would agree to initiate design studies for a larger 650-900 MW fast breeder reactor.

They would include a study of alternative configurations both of the uranium/plutonium cycle and the thorium/uranium cycle.

When the DoE's authorisation bill came before the full Committee on Science and Technology last week, Representative Walter Flowers (Dem.-Ala.), chairman of the subcommittee which had previously recommended funding for the Clinch River project, offered an amendment which was broadly in line with the administration's proposed compromise.

However in spite of the support of both those committee members who agreed with the president's desire to terminate the project, and others—including the chairman, Olin Teague—who although supporting the project, were prepared to accept the compromise as the best solution that could be expected in the circumstances, the amendment was defeated by the close margin of 20 votes to 18.

In its place the committee accepted by 25 votes to 14 an alternative amendment offered by Mrs Marilyn L. Lloyd who represents the district of Tennessee in which the breeder would be built, authorising the expenditure of \$35 million on the new study proposed by the Department of Energy, but leaving standing the subcommittee's proposal that work on Clinch River be continued.

In the long run, the committee's refusal to authorise funds to terminate the project is unlikely to have much effect. Dr Schlesinger made it plain, in offering his compromise, that the administration considers Clinch River is as good as dead; and in spite of the recent publicity given to the Comptroller General's apparent statement in a

letter that Carter's efforts to terminate the project were illegal, the president's supporters point out that elsewhere in the letter it was admitted that provisions in the contract did exist for it to be terminated unilaterally, although this could involve drawn-out litigation.

Although the president has been consistent in his opposition to the CRFBR—which has already involved the joint investment by Congress and private utility companies of about \$800 million and has reached the components testing stage—his argument has shifted.

Initially Carter's main, and widely-publicised, argument was the threat of the further proliferation of nuclear weapons implied by the use of plutonium in the fast breeder cycle.

Over the past year, however, partly because of a lack of success in convincing other countries of the logic of the US stand, and partly because of internal pressures and criticism, the proliferation argument has figured less and less prominently. In its place the debate has shifted to the size and shape of future energy demand.

The utility companies and the nuclear construction industry, following carefully designed plans for continuous expansion in which much capital has been invested in the promise of eventual returns, have argued that in order to meet projected energy needs, commercial orders for fast breeders must be placed by 1990 at latest. Clinch River, as both argued with a single voice before the energy subcommittee two weeks ago, is the necessary next step in the "orderly development" of a fast breeder programme.

President Carter, on the other hand, armed with declining projections of future population, the development of alternative energy technologies, and other leaves taken from the "soft technologists'" book, has, with the Department of Energy, been able to argue a case that fast breeders will not be required before 2200.

It is the space between the forecasts

Below: artist's impression of Clinch River



of the nuclear establishment and its critics that has provided Carter with room for political manoeuvring. And to give in at this point on the Clinch River issue would do little to help a presidential image already tarnished by an apparently equivocal stand on issues such as the neutron bomb.

The compromise solution has, admittedly, not pleased the environmentalist lobby, who point out that it merely reinforces a commitment to a nuclear future, and that it is, in the words of James M. Cunie of New Directions, "a *de facto* re-establishment of a commitment to move forward with commercial demonstration of the breeder reactor."

The environmentalists, having previously supported Carter's efforts to terminate Clinch River, have urged that the compromise be rejected by Congress—which, if nothing else, will make it easier for the nuclear supporters who oppose the environmentalists to accept it. But even the science committee vote, plus the likelihood that the appropriations committees in both the House and the Senate will continue to support the project, does not make a presidential veto inevitable.

Two alternatives remain. One is that Democrats in Congress, aware that there is little chance of raising the two-thirds majority needed to over-ride a

presidential veto, and unwilling to embarrass Carter—particularly in an election year—by further confrontation, will introduce an amendment along the lines of that proposed by Mr Flowers in debate on the DoE bill.

The other possibility is that the Senate Committee on Energy and Natural Resources, which meets soon to decide its position on the president's budgetary request, will accept the Schelsinger compromise, thus opening the way for it to be debated between Senate and House in conference.

If all else fails, and both houses reject the compromise formula, another presidential veto seems inevitable.

David Dickson

Little renewable energy until next century

RENEWABLE sources of energy—wind, wave, sun, geothermal heat and the like—are, you may be surprised to learn, taken seriously at the United Kingdom Atomic Energy Authority, Harwell. Dr Freddie Clarke, Research Director (Energy), who is responsible for the non-nuclear energy programmes of the UKAEA, delivered his assessment of alternative energy futures—broadly a thumbs down until next century—to a one-day meeting on energy in London last week. The meeting was organised by the National Federation of Women's Institutes, and was opened by the Prime Minister.

What could we consider to be a "substantial" contribution to UK energy supply by the year 2000? We now used 330 million tonnes of coal equivalent, Dr Clarke said, of which some 200 mtce were in oil and gas—forms which might be scarce and more expensive "sometime after 2000".

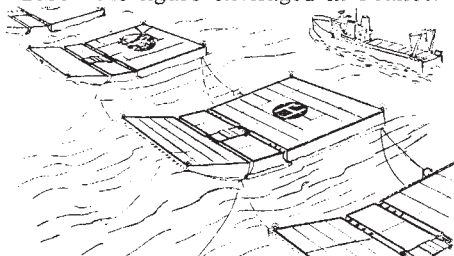
Official forecasts put demand at 500 mtce by 2000. The possibilities for meeting this increase, Dr Clarke said, are conservation, coal, nuclear, and renewable sources. But no contribution of less than 50 mtce per year could be considered substantial, he thought, and he set out to measure the renewable potential against this yardstick.

The conclusion he reached was that by 2000 there would only be a "minor" contribution from renewable sources of energy. But "somewhere in the world" they are economical here and now. Only if there were no technical setbacks in the development of renewable sources (an unlikely event), might we reach take-off—application in the UK on a substantial scale.

Some 30% of the UK's energy is for low temperature space heating, which makes geothermal heat look a good bet. The temperature of rocks increases at about 30 °C per km, and some deep sedimentary rocks are

penetrated by water, which can be drawn up to provide heating. A block of flats in Paris already receives 70% of its heat from such a source. In the UK, the midland valley of Scotland, the East Yorkshire-Lincolnshire basin, and the Hampshire basin are the best sources.

But this form of energy is not easily transported, and we have to see if demand is where the water is. The match in the UK is not very good, and Clarke could see little more than 2 to 3 mtce from this source by 2000—the figure envisaged in France.



The Cockerell raft: "well into the next century" before wavepower makes a substantial contribution

The "hot rock" technique, on the other hand, in which dry rocks are fragmented at 2 to 3 km depth and cold water pumped down to heat up, could yield several times the 2 to 3 mtce figure. Early results from a test site at Los Alamos "look encouraging" said Clarke.

Household solar energy, used for water heating, cost £600 to £700 to fit. At a current return of £20 to £40 per year, it was not yet economical—though the Government was researching ways to reduce the capital cost. It might pay by the 1980s. But here was a problem for all energy accounting: we don't know how easily the new technologies can be developed. Also, solar energy came mainly in the summer, and we need our heat mostly in the winter. Yet it was

possible, said Clarke, if everything went well, that solar energy could contribute 50 to 100 mtce by 2000.

The remaining renewable sources would be most likely to generate electricity rather than direct heat. A 50-metre windmill (about as high as an electricity pylon) could be economical now. But the interaction of an erratically turning mill and the national supply grid—into which it would feed—might cause problems. And the most efficient land-based sites for the mills would be on coastal hills in some of the most unspoiled parts of the country. The mills could be based off-shore, but would be more difficult to engineer there. On-shore, using the best sites, they could provide 6 mtce per year, Dr Clarke estimated. At sea this might rise to 20 mtce per year.

Wave power is the "most promising" renewable source in government eyes, with 100 mtce per year by 2000 "not inconceivable". But this system faces the greatest engineering difficulties.

The Department of Energy began a study of wave power. "Something struck me greatly when I saw artists' impressions of the full-scale machines" said Clarke. "It's the scale of the things". Clarke showed pictures of projects such as the "Salter ducks" and the "Cockerell rafts". "One element of these devices could contain this hall (the Central Hall, Westminster)" said Clarke. And they must run reliably even in the roughest part of the year. They represented "a new dimension in engineering" which was going to take time. It would be well into the next century before they make a big contribution, said Clarke.

Finally, tidal power, even using the best sites in the UK, can provide no more than 10 mtce per year and so cannot be a substantial contributor. It is also at present uneconomic.

Robert Walgate

Particle physicist to become a priest

JOHN Polkinghorne has become a minor celebrity lately—physicist turned priest: who'd have thought of that these days? "Prof takes the cloth" headlined the *Observer*. The local paper had called. The church papers were full of him. But he sees no conflict. "Science and theology are not not so very different" he said.

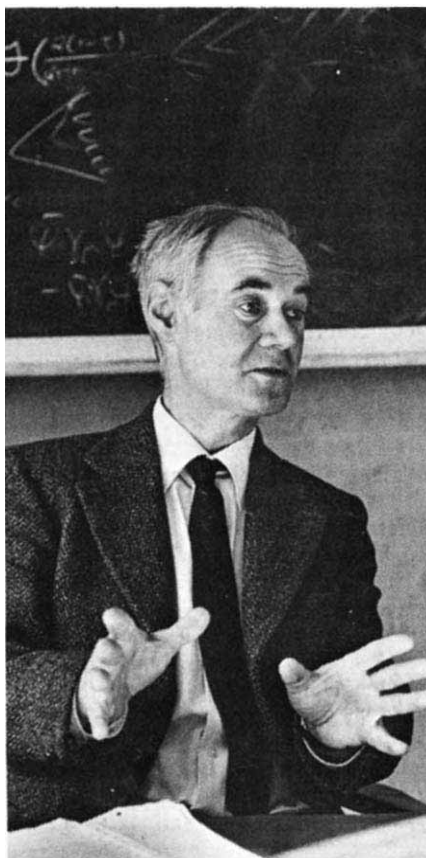
Not that Polkinghorne is going to become a theologian. "I doubt that I could become a professional theologian; there are a lot of skills you have to acquire and I'm starting too late to do that." He wants to learn as much theology as he can and is beginning to teach himself Greek; but the job he would really like is a teaching ministry—"I enjoy trying to explain things to people".

Polkinghorne, 48 years old with 25 years in the subject, explains why he is leaving particle physics. "I've had a good run for my money for 25 years. I've played a small but useful role. But the subject is continually changing and enriching itself." (An allusion here to the new wave of interest in "field theory" in particle physics—a subject he has been involved in only obliquely).

"I admire the way my younger colleagues can move into this sort of game. Partly through time but more through a question of mental flexibility, I don't think I can make that sort of move. I could therefore tick over for a few more years, but I don't think that would be a tremendously useful use of my time. I've done most of what I can do. If I felt I was going to stick in it and make reasonably useful contributions I would probably stay with it, and regard it as being a vocation to do so."

"When you're young you like to learn new things—that's part of the fun—but if I'm honest it gets slightly wearisome as you get older, particularly in a subject like particle physics where genuinely new things do come along at quite a surprising rate."

Polkinghorne does not think the techniques he has worked with, dealing



Robert Walgate talks to John Polkinghorne (above) who resigns his chair in physics at Cambridge next year

with "scattering amplitudes" directly rather than any underlying fields, are now useless—"they will continue to be applied in different contexts". He is writing a book for graduate students to "hand on a bit of the know how" and that will be "as useful a role as I can play".

"I don't think I'm exceptional in this respect. Most people in my subject do their best work early, and they tend to do the same tricks later on. There are some subjects in which these tricks remain very important,

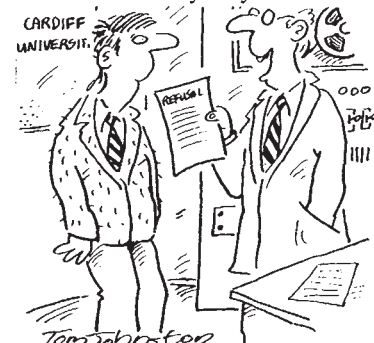
subjects whose basis is well understood, but elementary particle physics, relativistic quantum mechanics, is continually changing. There's less room for the wise old man. I think his wisdom is limited."

He also feels that, for good or ill, he is occupying a sensitive and important position in particle physics. A lot of talent passes through Cambridge, and whoever holds his chair has a significant influence over the way that talent turns. "That makes me more critical of what I can offer."

Altruism? "It's not entirely altruistic. I shall do less well as time goes on—and that would be less satisfying to me. I need to feel that I'm doing the right job. It would be nice if, having gone up in the academic system, one could come down again. I could earn my keep as a lecturer somewhere or other without getting hypercritical about it."

Would he recommend others to follow his course? He chuckled. "That's not for me to do. It varies so much from person to person and job to job. But it's a general phenomenon that you get less active and less in touch with the real developing point of the subject as you grow older. But there are many ways of dealing with it. Not everyone wants to become a priest!"

Polkinghorne decided two or three years ago to make a change; and he did not take long to decide that his change should be to the priesthood. "I've been a confessing Christian all my life". He resigns in October 1979, and then has a three-year teaching fellowship at Trinity College ("that will be a considerable financial help"). He will spend three more years in Cambridge to see his younger son through school. Two of these years he will spend at a Cambridge theological college. "I've a lot to learn." His third year he hopes to spend as a deacon, as an honorary curate in a parish near Cambridge; and then he will be a priest concentrating on the mystery of life rather than the mystery of matter. □



"Maybe if we told Callaghan that we're doing research into evidence of Labour voters in interstellar space. . . ."

Scientists appeal to PM over grant

SIR FRED HOYLE and Professor Chandra Wickramasinghe, of University College, Cardiff, Wales have taken the unusual step of writing to the Prime Minister, Mr Callaghan, about the rejection of their grant application for further research on polysaccharides in interstellar space. Their grant application of last September to the Science Research Council for £7,500 to buy a mini-computer was turned down by the council's astronomy committee on the

basis of insufficient merit. It seems to have been handled in the normal way by being sent out to referees, although referees' reports were not transmitted to the applicants.

Mr Callaghan has as yet shown no enthusiasm to intervene in what is undoubtedly a highly controversial field, leading as it does to speculations that the origin of life on Earth involved the assembly of material which exists widely in the interstellar medium. □

DNA: the controversy continues

● Milan: regrets at WHO over ban on research

● Washington: America ponders legislation (overleaf)

WHO looks for benefits from genetic engineering

Peter Newmark reports on a recent European meeting on the application of DNA

THE WORLD Health Organisation, in a symposium held last month at the Giovanni Lorenzini Foundation in Milan, showed its interest in the "scientific developments and practical applications" (the symposium subtitle) of genetic engineering. This is a sign of the times, for WHO's original concern was the danger of genetic engineering: the potential spread of disease carried by newly created micro-organisms.

Four years on, the dangers seem less. Hazards are now spoken of as 'conjectural' not 'potential'. Offices and officers are not concerned with 'bio-hazards but with biosafety'. One must now accentuate the positive.

The negative began in 1974 shortly after new techniques opened the door to the experimental incorporation of foreign genes into those of bacteria (forming "recombinant DNA"). Cloning of the recombinant DNA by culturing the bacteria promised both academic and commercial rewards.

This was excellent news for those who wished to study the molecular structure, and ultimately function, of the genes of higher organisms, up to and including man. Better still for some was the possibility that bacteria into which had been incorporated the gene for, say insulin, would without hesitation decode the instructions and synthesise the hormone, at a commercial rate.

The darker side of such experiments, as envisaged in 1974, turned on the likely bacterial host for the foreign genes. It was the bacterium *Escherichia*

coli (*E. coli*), some strains of which (but not the laboratory kind) are pathogenic to man. Many strains of *E. coli*, can colonise the human gut. What if a laboratory *E. coli*, engineered to synthesise insulin, escaped, found its way into guts and there poured out insulin? Worse still, what if the gene that codes for the production of cholera toxin was incorporated in the *E. coli* that escaped?

Worried by these possibilities a group of scientists, including pioneers of the recombinant DNA techniques suggested a moratorium on experiments until precautions could be designed to match the supposed hazards. The moratorium triggered off an avalanche of meetings, concern, guidelines, regulations and bureaucracy.

It is hardly surprising that many of the speakers in Milan came to honour the pioneers but certainly not to praise them. Time and again regret was expressed that the moratorium had ever been called. Demands for a public recantation by the whole group were muted by the thought that ailing public confidence in scientists would hardly be resuscitated by the spectacle of erstwhile experts admitting that they had cried wolf.

Why have opinions changed so drastically in less than four years? The strongest view at the meeting was that the original worry came from a close-knit group of molecular biologists who had not taken enough trouble to consult experts in infectious diseases, medical microbiology and epidemiology. Furthermore evidence has been

accumulated since 1974 to suggest that the laboratory strain of *E. coli* could not survive in the wild and that in any case the process of incorporating foreign into bacterial DNA is a naturally occurring one. The new evidence, however, does not seem substantial; those at Milan witnessed some unseemly clutching at straws.

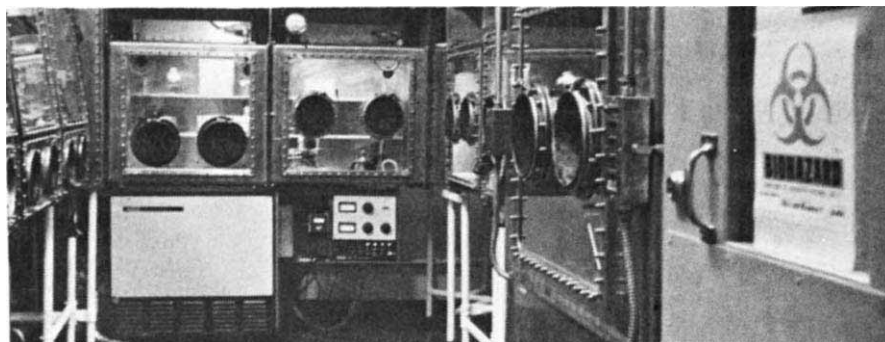
Meanwhile regulations or guidelines have placed a variety of temporary barriers in the way of research. These have been met through both physical and biological containment. The former has involved building or equipping laboratories to strict standards. The strictest of these have just been met for the first time in a laboratory,—the Frederick Cancer Research Centre at Fort Detrick in Maryland. That will be followed, probably in six months time, by a high containment facility at the European Molecular Biology Laboratory in Heidelberg. Biological containment has been achieved by the "construction" of enfeebled vectors (the plasmids or phage that are used to carry the foreign DNA) or bacterial hosts. Most successful have been the American version of *E. coli* patriotically named χ 1776 and the derivative of λ phage optimistically named Charon. In Milan the question arose whether either were vigorous enough to be really satisfactory. Perhaps the British version devised in S. Brenner's laboratory in Cambridge and currently under test, will be better.

Another way to avoid the barriers surrounding the use of *E. coli* is to replace it by a host bacterium that is not pathogenic. Three different speakers in Milan described the development of *Bacillus subtilis* as an alternative host. Even if *E. coli* is as safe as some would claim, the availability of an alternative is going to be valuable since it is already clear that cloned genes may be expressed in one species but not in another.

Guidelines now exist in some form in every country which carries out recombinant DNA research. There are discrepancies between countries which sometimes cause professional jealousy. But J. Tooze, executive secretary of EMBO, suggested that harmonising guidelines internationally would be catastrophic, creating a stranglehold and preventing change.

The direction of change desired by all speakers in Milan was towards fewer restrictions. Already the French have achieved that. Their national regulations, adopted quietly three months

Inside Fort Detrick, Maryland: built to the strictest safety standards



ago, are an order of magnitude less restrictive than the National Institutes of Health (NIH) regulations. Tooze hoped that Italy and Japan, still finalising their regulations, would follow France.

Several representatives of large industrial concerns were optimistic regarding the potential benefits of recombinant DNA technology to health and industry. The first benefit is likely to be improved bacterial strains, some providing increased yields of enzymes of industrial importance. Further ahead should lie the production of highly specific vaccines. For example, S. Falkow from Washing-

ton University, Seattle, spoke of the potential as a cholera vaccine of a product of the cloned gene for the heat-labile *E. coli* enterotoxin.

The most tangible of the much-vaunted benefits of recombinant DNA, however, lies in the production of hormones. As far as can be judged the running in that race is being made by smaller industries spawned by the technology and with very close links to universities. In particular there is Genentech, a firm whose founders include H. Boyer of the University of California, San Francisco. His university group has recently succeeded in chemically synthesising

the gene for the hormone somatostatin incorporating it into a bacterial plasmid and then showing that when the plasmid entered host *E. coli*, the bacteria synthesised the hormone. The same techniques are now being applied to the production of insulin, a remarkable goal that is on target to being accomplished before the end of the year. Boyer estimates that the cost of insulin produced in that fashion should be no more than for that now purified commercially from slaughterhouse pancreatic tissue. Furthermore the insulin from Boyer's technique should be purer and could be of human origin and sequence. □

Friends of DNA fight back

David Dickson describes an unpublicised campaign by university administrators to halt community participation in DNA legislation

FOUR American universities have employed a Washington attorney as a professional lobbyist—although initially unregistered—in an attempt to prevent the passage of legislation allowing state or local participation in establishing regulations covering research using recombinant DNA techniques.

Already three of the four universities—Harvard, Stanford and Princeton—face the possibility of state legislation covering such research; and Harvard has been made subject to local ordinances passed last year by Cambridge City Council.

At the fourth, Washington University in St Louis, little research involving recombinant DNA is being done, but an official said last week that they were concerned with the principle.

The issue of Federal preemption—Federal legislation that overrides state or local initiatives—lies at the heart of the current dispute. Public interest groups argue that local communities should be directly involved in drawing up regulations for research carried out in their midst, pointing to events at Cambridge to show how this can be done in what they claim to be a reasonable and moderate manner.

In contrast, many of the scientists argue the need for Federal preemption to provide uniform safety standards and prevent research workers moving to localities with less stringent regulations. Dr Philip Handler, president of the National Academy of Sciences told a Congressional Committee last week that the Federal preemption in the Bill now before the House of Representatives was “a particularly constructive feature”.

Much of the opposition from the scientific community which led to last year's suspension of Congressional activity on DNA legislation was the result of a well-publicised campaign. At the same time, however, an unpublicised lobbying effort was conducted by university administrators concerned that DNA legislation could set a dangerous precedent for demands for community involvement in other areas of university activity.

Although the lobbying attempt has been done primarily through Harvard, more than 70 individuals at over 35 universities throughout the country have been involved through an informal network known as the ‘Friends of DNA’. Only half of these are scientists; the rest are involved in university administration, in particular through offices of public affairs or similar liaison units.

The universities’ concern at the implications of DNA legislation was first raised a year ago when Representative Paul G. Rogers, chairman of the Health and Environment Subcommittee of the House Committee on Interstate and Foreign Commerce, introduced an Administration-backed Bill which would have permitted state or local legislatures to impose conditions on DNA research more stringent than Federally-agreed guidelines determined by the National Institutes of Health.

Many scientists reacted strongly against this suggestion. On 26 April last year, a resolution signed by 13 members was passed by the National Academy of Sciences claiming that the proposed legislation would subject DNA research to “arbitrary regula-

tions” at the local level which would “severely degrade” the capability of biomedical research.

Dr Harlyn Halvorson of Brandeis University, then president of the American Society for Microbiology, started a campaign among professional scientific societies through the Inter-Society Council for Biology and Medicine whose objectives included “a single set of uniform national standards governing recombinant DNA activities”. At the same time, the four private universities—Harvard, Stanford, Princeton and Washington—agreed to contribute towards the expenses of a professional lobbyist, Ms Nan Nixon, who would argue for Federal preemption in any DNA legislation.

She has conducted a vigorous lobbying campaign on the universities’ behalf in collaboration with Mr Donald Moulton, until recently Harvard’s assistant vice-president for Community Affairs, now working in commercial real estate and acting for Harvard as a consultant. (Although Ms Nixon has provided the Congress with details of her activities—as professional lobbyists are required to do—from the beginning of July 1977, she did not in fact register until February 28 this year, coincidentally handing in her registration two days before an article describing her activities as a registered lobbyist appeared in the *Harvard Crimson*.)

The scientific and university-based lobbyists soon met success. Mr Rogers and the Administration revised their Bill to state that the Secretary for Health, Education and Welfare could permit local, more stringent legislation only if it could be demonstrated that this was “necessary” (a word which has since taken on an important significance) to protect the health of people and the environment.

This was the form in which the Bill went before the House Commerce Committee, under the chairmanship of Representative Harley Staggers, last

autumn, and it won from Ms Nixon the statement that she was "generally in favour".

However, during committee consideration of the Bill, an amendment was passed changing "necessary" to "reasonable". Shortly afterwards, the Harvard lobbyists withdrew their support, and Mr Staggers who had also been contacted by a number of scientists questioning the whole wisdom of the Federal legislation, blocked further discussion by the committee.

The universities' lobbyists then agreed with Mr Staggers to produce a new Bill which would include preemption of all Federal agencies (except the Occupational Health and Safety Administration) by the Department of Health, Education and Welfare, and Federal preemption of all state and local governments. A draft Bill containing these provisions was circulated to all members of the Friends of DNA network in the middle of November.

In a subsequent memorandum, Ms Nixon and Mr Moulton circulated a copy of the local ordinance covering DNA research which was being considered by the City Council of Berkeley, California—and has since also provided the model for legislation that has been proposed to the State of Massachusetts. The memorandum said the ordinance "clearly illustrates the need for Federal preemption".

A revised version of the Bill was prepared at the end of December with the aid of an Arlington lawyer, and a copy sent to Mr Staggers, who introduced it virtually unchanged into the House of Representatives on 19 January. Meanwhile, members of Mr Rogers' staff had produced their own revised Bill which omitted any discussion of the preemptive issue.

Following discussions, a joint Bill supported by both Mr Staggers and Mr Rogers was prepared and made public at the American Association for the Advancement of Science meeting in Washington in February. It contained the preemption provision as suggested by the universities' lobbyists.

In a memorandum from Ms Nixon and Mr Moulton, the Friends of DNA were told that "several of us have read the draft Bill and in its total context feel it's an acceptable Bill". However, members of the network were warned of a "major problem" that, particularly in the preemption section, "there is no padding, and therefore its acceptability may not continue if it is amended during mark-up".

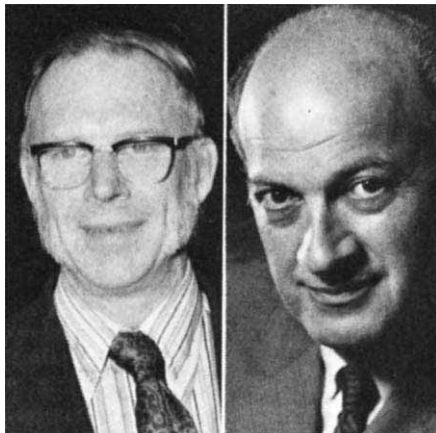
The memo continues: "If you agree with our opinion of the Bill and our concern for amendments, please make your opinions known to Mr Staggers, Mr Rogers, and all members of the full

committee".

At the mark-up session on 13 March, the full committee considered the joint Staggers-Rogers Bill, now formally known as HR 11192. Two amendments seeking to rephrase the preemption language were voted down. The first, which would have reversed the Bill by placing responsibility on the Secretary of HEW to show that local, more stringent regulations were *not* necessary, had been supported by pressure groups such as the National League of Cities; the second, which would have changed "necessary" to "reasonable", was rejected on a voice vote which the chairman, in a hotly-disputed decision, refused to put to a count.

Deep political differences are revealed by the committee's report on its mark-up session. The majority report states that "local action is not always based upon careful consideration or understanding of the available facts on a particular issue".

In contrast, the six dissenting members who voted against the final version



Science lobby: Halvorson (left), Handler

of the Bill state that "our fundamental concern as to the preemption issue is with the notion that a scientific elite and an insular federal bureaucracy know what is best for the people", an attitude which is "contrary to the principles of Jeffersonian democracy . . .".

In Cambridge, ex-mayor Alfred Vellucci, who was largely responsible for moves that led to the setting up of a local Biohazards Committee to oversee the regulation of DNA research at Harvard and Massachusetts Institute of Technology, called a press conference to declare that he would "put up a fight" against Federal preemption. Sheldon Krinsky of Tufts University, a member of the panel which had suggested setting up the committee, announced that "everything that occurred in Cambridge will really be meaningless if this legislation is passed".

Unless an amendment is offered when HR 11192 is debated on the floor, the Staggers-Rogers Bill seems likely to

prevail in the House of Representatives, with full Federal preemption. Mr Moulton said last week that he was "optimistic" about the outcome.

In the Senate, however, things are much less certain. Senator Edward Kennedy, the author of a Bill which was the focus of attention last year until he dropped his support for it, has been consistently opposed to Federal legislation. And last month, rather than introduce the Staggers-Rogers Bill into the Senate as had been predicted, the Senator introduced his own Bill in which preemption is not mentioned.

The universities' lobbyists have been trying hard to get this reversed. The Bill prepared for Mr Staggers was also given to the staff of Senator Jacob Javits of New York, who had previously argued against Kennedy, in favour of strong Federal preemption.

A few weeks ago, the chances of success looked high, since the various senators (apart from Kennedy) most closely involved with the DNA issue seemed to favour fairly strong Federal preemption. But two recent events are giving the lobbyists cause for concern.

The first is a subtle but significant shift in the Administration's position on preemption, following a heated dispute between the domestic counsel of the White House, who has in general been against strong Federal preemption, and the director of the National Institutes of Health, Dr Donald Frederickson, who has in general been in favour of it.

Presenting evidence to a subcommittee of the House Science and Technology Committee last week, Dr Gilbert Omenn, assistant director of the President's Office of Science and Technology, revealed that a compromise had been reached. He said that while the Administration supported HR 11192, it felt Congress's intentions could be met with a test that the change proposed by the states or localities be "reasonable" rather than "necessary".

A further factor likely to affect the Senate debate is the imminent publication of the report on hearings held by Senator Adlai Stevenson last November. These are expected to contain detailed recommendations for legislation, including suggestions for a weaker preemptive provision which could be put forward by Stevenson as a compromise between Kennedy and Javits.

Neither of these two initiatives has been welcomed by the Harvard-based lobbyists. Although the debate on the adequacy of the NIH's safety guidelines has calmed down in a way that most scientists find acceptable, the political reverberations threaten to linger on into what could again become a long, hot summer debate. □

correspondence

Israeli participation in Soviet conference

SIR,—It was with great disappointment that I read the letter (16 February, page 605) on Soviet 'discrimination' against Israeli scientists in connection with the International Meeting on Ferroelectricity held in Leningrad 18-23 September, 1977. I was one of the vice-chairmen on the conference organising committee. In this capacity and on the instruction of the organising committee, I assisted in arranging for Israeli scientists to participate in the conference.

In particular, I spoke twice to Dr W. Merz of RCA Zurich in July and informed him that the Israeli scientists' papers had been accepted and that visas would be issued. Later on in mid-August, on the request of the organising committee, I called Dr Merz from my laboratory in Moscow and consulted him on the best place from which to issue the visas to the Israeli scientists.

At the time four names were mentioned to me; unfortunately I can now only recall that of Mr S. Havlin. Dr Merz and I agreed that visas for two participants (Mr Havlin was one of them) would be issued at the Soviet Embassy in Rome. The other two would be issued in Vienna. I asked Dr Merz to inform the Israeli participants accordingly and was quite sure that the matter was settled.

Nevertheless, when I returned to Vienna on 12 September, I checked on whether or not the visas had been issued by the Soviet Embassy, and I was astonished to find that the Embassy had received no applications for visas from Israel. I now understand that the Soviet Embassy in Rome received no visa request either. Last November, I contacted Dr Shuvalov in Moscow to clarify some details of the arrangements made for the Israeli scientists. I understood from him that the organising committee was not informed that Mr Havlin had applied for his visa in Paris and not in Rome, until the beginning of September. Action had been taken immediately and on 9 September the Soviet Embassy in Paris was instructed to issue the visa.

It is regrettable that Israeli scientists were not able to participate in the International Meeting on Ferroelectricity. It is regrettable that through no fault of ours but because of understandable difficulties in communication between Israel and the Soviet Union details of arrangements for conference participation were delayed. Considering the problem of Israeli participation in the conference, one has to take into account in addition the fact that for purely technical reasons the distribution of the information circular was delayed to all participants.

I find the last part of the letter completely inappropriate. Rather than avoiding difficulties in future, it would be more likely to create new ones. The recommendations to impose certain conditions when arranging international conferences in the Soviet Union are based

on a personal understanding of a few international problems and neglect the tremendous international advantages in cooperation between West and East. On the basis of the experience gained through the Leningrad conference, the organising committee could, but never would, come to some completely opposite conclusions. For example, we could judge the fact that just two scientists tried to participate instead of the four we were informed of, as an additional difficulty created intentionally. Further, the appeal in the letter to impose all those restrictions before the matter is clarified could be converted—but will not be—to an appeal to the relevant scientific community not to invite Israeli scientists to international conferences until the writers apologise for the distribution of improper information.

I. S. ZHELUDEV

Moscow, USSR

IQ and environmental variables

SIR,—Stephen Gould is quoted (23 February, page 699) as saying that Arthur Jensen's work rests on "a fundamental fallacy", namely, that he fails to understand that an average difference between populations in IQ need not be genetic "even if the IQ is heritable within the black and white populations". The scientific assertion is correct, but it is incorrect to attribute this reasoning to Jensen.

What Jensen does argue is that there is a significant connection between 'within group' and 'without group' heritability such that high within group heritability lessens the probability of environmental explanations for a difference between groups. And he is right. We can see why this is so simply by imagining the environment as made up of a large number of elemental factor variations which may or may not affect IQ, and then rephrasing the question as follows: would any of these factors be excluded as a cause of an average difference between groups if the within group heritability were (say, for expository purposes) 1.0? The answer is yes: namely, that class of environmental factors where the difference (if any)

between the two groups is within the range of variation already sampled in the group in which a heritability of 1.0 was determined.

We can illuminate this by developing the example cited by *Nature*: assume two groups, A and B, differing in average height, and let us ask if dietary factors could account for this difference. If whatever group difference that existed in the diets of A and B were within the range of diet variation already examined in determining the heritability of height in A, then a heritability of 1.0 in A certainly would exclude that diet difference as a cause of the group difference in height. If on the other hand the heritability were zero (and especially if this were a result of diet variation in A), then any significant difference in diet between A and B would be a plausible cause of the height inequality.

Thus, by excluding a sizable class of environmental variables, high within group heritability unarguably does raise the probability of the remaining classes of possible explanations, of which there are two: either a genetic difference, or else an environmental difference between groups that is *not* represented within groups (and hence was not sampled in the determination of within group heritability). Since it happens that many of the environmental correlates of racialism have precisely this kind of distribution (segregated schooling, psychological injuries associated with being a minority, and many more), this latter class is by no means an empty formalism. Nevertheless, whether some member of this rather than the genetic category is the cause of the group IQ difference needs demonstrating, not merely hypothesising. It is in evaluating the relative likelihood of these two formal possibilities that an assessment of programmes like Head Start (that explicitly predicated an environmental deficit as the origin of individual and group differences in IQ) becomes relevant. And it is primarily the unimpressive success of this and many similar "environmentalist" prescriptions that led Jensen to prefer a genetic explanation.

That is, it is not, *pace* Mr Gould, an elementary misunderstanding of heritability theory at all that is the *fons vitae* of Jensen's position but rather the consideration over and above theory of a vast body of empirical evidence. Thus those among us who would dispute Jensen cannot confine their criticism merely to abstract theorising; what is needed to 'falsify' the genetic interpretation once and for all, and simultaneously to provide the clue that could lead to sound and compassionate social policy, is the empirical demonstration of at least one extant environmental difference between racial groups that, when equalised, significantly reduces the IQ difference. Until this can be done, the 'environmentalist' interpretation is unlikely to contribute much of a concrete nature to the shaping of wise social policy.

WILLIAM R. HAVENDER
Berkeley, California, USA

Sorry, for copyright reasons some images on this page may not be available online

The Times

Arthur Jensen

news and views

Dodging Heisenberg

from N. MacDonald

THE ultimate limit to the precision of physical measurements is set by quantum considerations. Any amplifier is noisy because of zero-point energy. All mechanical measurements, for example on the massive oscillators used in attempts to detect gravitational waves, are limited in precision in so far as they involve simultaneous measurement of position and momentum and so run up against the Heisenberg uncertainty principle

$$\Delta x \Delta p \gtrsim \frac{1}{2} \hbar$$

In a recent paper K. S. Thorne, R. W. P. Drever, C. M. Caves, M. Zimmermann and V. D. Sandberg (*Phys. Rev. Lett.* **40**, 667; 1978) propose a radical new approach to this limitation. One variant of this approach has simultaneously been advocated by a group in Moscow (Braginsky, Vorontsov & Halili *J.E. T.P. Lett.* March; 1978).

Both these groups are engaged in planning experiments for the detection of gravitational waves. Both the first generation of gravitational wave detectors, and those at present under construction, are very far from the quantum limit of precision, and also very far from the precision at present considered necessary if one is to be able to detect several gravitational wave pulses per year. The significance of the quantum limit in this context was first pointed out by Braginsky. The basic quantities needed to assess the precision of measurements on the oscillation of a bar of mass M , length L , are the duration T of the pulse one hopes to detect, and the fractional strain $\delta L/L$ caused by this pulse. T is thought to be of the order of 10^{-3} s, while pulses originating from supernovae out as far as the Virgo cluster of galaxies, and likely to arrive at the rate of three each year, are thought likely to yield $\delta L/L$ about 10^{-20} to 10^{-21} .

Approximating the oscillating bar by a mass on a spring one is, in effect, measuring the extension L and the velocity $\delta L/T$. The uncertainty principle thus leads to an error Δ in the strain given by

$$\Delta^2 \gtrsim (\Delta L)^2 + \left(\frac{1}{2} \frac{\hbar}{\Delta L M} T\right)^2$$

which leads to the estimate

$$\Delta \gtrsim \left(\frac{\hbar T}{M}\right)^{\frac{1}{2}}$$

For a mass of 10 kg this is about 10^{-19} , and it falls only as the inverse square root of the mass for more massive oscillators. So even if other sources of imprecision are removed from the planned third generation of detectors, so that detection of the faint pulses is in prospect, some way round this quantum limit must be found.

Braginsky proposed a few years ago to circumvent this difficulty by means of direct measurement of the energy of the oscillator, that is to say of the number of quanta (phonons) in the excitation of the bar. This approach encounters several difficulties. For example to count the number n of quanta without creating or destroying any quanta requires that the interaction Hamiltonian H , expressing the coupling of the oscillator and the detector, must commute with the number operator N , which means that H must involve the squares of the coordinate and momentum of the oscillator displacement. With such minute displacements this would be very hard to achieve.

The new proposal is compatible with linear coupling of the oscillator and the detector. In terms of the complex amplitude

$$X_1 + iX_2 = (X + iP/M)e^{i\omega t}$$

it is apparent that an alternative pair of conjugate variables to the usual X and P are X_1 and X_2 . Any device requiring measurement of both X_1 and X_2 runs up against the uncertainty principle. However it was noted by Thorne that it is possible to get all the information needed about the passage of a gravitational wave pulse from measurements of X_1 alone.

Since the best way to bring out the physical implications of the uncertainty principle is by means of 'thought experiments', it may be helpful to think about the stroboscopic illumination of an ideal pendulum. Such a stroboscopic experiment is mentioned in the paper by Thorne *et al.*, and this aspect of the problem is discussed in more detail by Braginsky *et al.* The pendulum bob is illuminated by a flash of light and its instantaneous position x thus determined. But this precise measurement is paid for by delivery of an arbitrary recoil to the bob from the scattered photon. The bob acquires unknown momentum, and so the future position is, in general, unpredictable. However, if one illuminates only at intervals of half the period of the pendulum, one can predict that values $-x, x, -x, \dots$ will be obtained. This amounts to measuring the eigenvalue of X_1 taking the time t of the first flash as one at which $\sin(\omega t) = 0$, so that $X_1 = \pm X$ at each instant at which the bob is illuminated, and the eigenvalue $x_1 = x$.

Continuous measurement of X_1 requires modulation of the output of position and momentum transducers by $\sin(\omega t)$ and $\cos(\omega t)$, without the intervention of any new source of noise. For example Thorne *et al.* discuss in some detail a torsion balance device for X_1 measurements for an electrical oscillator. There are severe difficulties in the way of a practical realisation of these measurements, but the principle of the new approach is applicable, not only for mechanical detectors of gravitational waves, but for detectors of any weak but classical disturbance. $\square$

More about the beam dump

from David J. Miller

PRELIMINARY results of the CERN neutrino beam dump experiment were discussed in a News and Views report some weeks ago (272, 205; 1978). Now that the papers of the three experimental groups are published (*Phys. Lett.* **74B**, 134, 139, 143; 1978) it is possible to examine the agreements and differences between them more closely.

They all agree on one new effect; the rate of production of events with an electron and no muon in the final state, compared with the rate for muon production, is much too high to be accounted for by the 'old fashioned' particles, the pion and the kaon. Such particles prefer to produce muon neutrinos which in turn prefer to give muons in the final state when the neutrinos interact. The beam dump experiment uses a special massive copper target in which the pions and kaons from high energy proton reactions are rapidly absorbed, instead of being allowed to decay and produce neutrinos as they do in normal neutrino experiments. To account for the high rate of electron events a 'prompt' neutrino source has been postulated, some short-lived particle which gives roughly equal numbers of all four kinds of neutrino—muon and electron, particle and antiparticle. The bubble chamber collaborations (the BEBC collaboration of Aachen-Bonn-CERN-London, Imperial College-Oxford-Saclay (Bosetti *et al.* *Phys. Lett. op. cit.*, 143), and the Gargamelle Milan-Orsay-Strasbourg-London, University College (Alibrán *et al.* *Phys. Lett. op. cit.*, 134) suggest that the production rate for this new prompt source should be around one in 200 proton interactions. The CERN-Dortmund-Heidelberg-Saclay (CDHS) counter collaboration (Hansl *et al.* *Phys. Lett. op. cit.*, 139) claim that they can explain this observed rate of electron events if prompt neutrinos are produced in around 1 in 1,000 proton interactions. It is not yet clear where the difference comes from. It may be in the beam flux calculations or in the model used to calculate the way in which heavy parent particles would be produced by the proton beam. At the rate seen by CDHS it is just possible that the prompt parents could be the charmed D mesons, although some special arguments are needed to explain why D decays have not been seen in experiments which looked directly for them.

CDHS group also sees an excess of positive muon events over the expected

number. This fits in well with the idea of a prompt source giving equal numbers of all four types of neutrinos. Pions and kaons from proton interactions produce many more neutrinos than antineutrinos. This is because both the proton beam and nuclei in the upper target have positive electric charge, so positive pions and kaons are more readily produced and these decay largely to positive muons accompanied by neutrinos. Negative pions and kaons give negative muons with antineutrinos. When neutrinos interact in charged current processes they produce negative muons (particles) and when antineutrinos interact they give positive muons (antiparticles, c.f. the positron, the anti-electron). If D mesons are the prompt source then they are probably produced as particle-antiparticle pairs, conserving the charm quantum number in the strong production process, and giving equal numbers of neutrinos and antineutrinos in their decays. So the CDHS data are again consistent with the D meson.

Trimuon events were the trigger for doing the beam dump experiment. In one of their previous runs with a normal neutrino beam CDHS saw 13 events with three final state muons with a total of 2.3×10^{17} protons on the target at 400 GeV/c. It was suggested that these might be caused by a new type of neutrino which would be produced promptly. In the beam dump experiment, with 4.3×10^{17} protons on target at 400 GeV/c, no trimuons were seen, so it is established that the trimuons come from the interactions of normally produced neutrinos.

All three collaborations have been able to place limits on the production and interaction of axions. These postulated neutral particles, with 'semi-weak' interactions and a very low mass were invented in order to solve a problem on the construction of gauge-field theories with parity and time-reversal invariance (Weinberg *Phys. Rev. Lett.* **40**, 223; 1978). The beam dump experiment is a good place to look for them, since they are supposed to be produced directly in proton collisions, not through an intermediate pion or kaon decay, and their semiweak interaction should allow many of them to pass through the thick shield of steel and rock which absorbs charged particles from the neutrino beam. The BEBC beam dump results show that axion production must be less than 1/30 of the predicted rate. This is in agreement both with the Gargamelle result and with a recent paper from an old Gargamelle beam dump experiment,

reanalysed for the axion by a group in Milan (Milan University preprint; Bellotti, Fiorini and Zanotti). The CDHS collaboration have larger statistics than these bubble chamber experiments. Their limit on axion production and interaction is a factor of 200 smaller than the predicted rate for one class of axion interaction, and 20,000 times smaller than predicted for another class of interaction. Unless its inventors can adjust their theories, the axion seems to be dead. □

Plasma on the wall at Culham

from M. W. Thompson

THE new mood of optimism in the plasma fusion business and the decision to site JET, the Joint European Torus experiment, at the Culham Laboratory of the UK Atomic Energy Authority, brought a sense of purpose to the Third International Conference on Plasma Surface Interactions held there recently*. If power-producing fusion reactors become a reality next century, solutions will have been found to the problems of radiation damage in the reactor structure.

Outside the fusion reactor vessel there will be neutron and gamma radiation similar to that outside the core of a fast fission reactor, for which the materials problems are already solved, although the upward shift of neutron energy from 2 MeV to 14 MeV, and the different intensities, will move us on to less familiar ground. But the real problems will be inside the reactor vessel. If, as many people expect, fusion reactors confine their plasma within electromagnetic fields, the vessel will be a vacuum chamber and its inner walls will be bombarded by almost every conceivable kind of radiation. Electrons, ions and neutral atoms of hydrogen and helium, a few heavier ions and atoms and electromagnetic radiation from visible light through X rays to gamma rays are expected. These particle energies will spread all the way from a few eV to several MeV. Besides these there will be the 14 MeV neutrons from the fusion reaction. In the current experimental systems only the neutrons and gamma rays are missing. Although the containment fields are intended to confine the ions and

* Held on 3-7 April. The proceedings will be published by North Holland as a special issue of the *Journal of Nuclear Materials*.

David J. Miller is in the Department of Physics and Astronomy, University College London.

electrons within the plasma during the fusion reaction, some will escape and it will be even more difficult to prevent energetic neutral atoms bombarding the walls.

At present it seems that the effects of the light ions and atoms present the most serious problem. In the first place they cause erosion of the walls by sputtering; they also become implanted in the walls and can blister and flake the surface. Release of atoms from the wall by sputtering is a serious problem because the presence of elements heavier than hydrogen in the plasma would increase the rate at which it radiates energy away as light or X rays. The plasma then becomes cooler and fusion more difficult. Solutions to this problem are on the one hand to reduce the sputtering of the walls by the choice of appropriate materials, and on the other to clean the heavy impurities produced by sputtering out of the plasma, or at least to prevent them reaching the hot central region where fusion is going on.

Sputtering of metals by light ions has been intensively studied for many years now, and for ions above 10 keV the main characteristics such as the yield of sputtered atoms and their angular distribution as a function of particle energy are reasonably well known. Below these energies, data are scanty and theories unproved, but it may be that at the lower energies the majority of bombarding particles will strike the wall. This possibility will clearly receive more attention from the sputtering fraternity, as will the difficult measurement of the energy distribution of sputtered atoms. Increasingly they will be thinking of peculiar effects such as chemical or reactive sputtering, which has emerged again after a few years out of fashion.

The processes responsible for the breaking up of the surface of metals when atoms of hydrogen and helium are energetically implanted into them are now becoming clear. There is a complex interplay of lattice defects and impurity atoms which can, in certain conditions of temperature and bombardment, cause gas-filled bubbles to form in the wall. It is thought that the stresses set up in the surface by radiation damage make the bubbles join up to form blisters which burst and form a flaky surface. Parts of the story are very familiar to those who have been in the nuclear materials game before. Fuel element swelling in fission reactors involves a similar set of processes.

Some unwelcome fluctuations in vacuum pressure have been turning up in plasma experiments and they seem to have their origin in the plasma-

Intermolecular forces

from Graham Richards

To activists in the field, the subject of 'intermolecular' forces is limited to the study of interatomic potentials and more specifically the interaction between rare gas atoms. Knowledge of the intermolecular potentials would allow bulk properties of gases such as viscosity, thermal conductivity and diffusion behaviour to be simply calculated; at present this process is essentially reversed and the determination of these bulk properties is necessary in order to calculate the intermolecular potential.

Although the repulsive part of the potential can be calculated from accurate *ab initio* Hartree-Fock calculations, reasonably direct measurements which lead to microscopic potentials were not available until the observation of dimers by spectroscopic means.

The dimeric species, Ar_2 , was first observed by Tamaka and Yoshino (*J. chem. Phys.* **53**, 2012; 1970) and the electric absorption spectrum later photographed at a dispersion sufficiently high to permit rotational analysis by Colbourn and Douglas (*J. chem. Phys.*, **65**, 1741; 1976). From a knowledge of the energy levels of the dimeric molecule the potential curve which gives rise to those levels can be deduced using the well-known Rydberg-Klein-Rees method.

Now Aziz and Chen have combined all the previous data to produce a simple precise intermolecular potential for argon (*J. chem. Phys.* **67**, 5719; 1977). The new potential fits both the accurate spectroscopic data and the high temperature viscosity values of Maitland and Smith (*J. chem. Eng. Data* **17**, 150; 1972).

surface interaction. It seems to be a complex case of implantation and sputtering resulting in the release of gases and vapours (for example, N_2 , O_2 , CO , H_2O) from the walls. Various schemes are in train to 'condition' the walls by bringing them in contact with low temperature plasma to release the gases beforehand. But because of implantation the walls can develop an embarrassing memory, releasing impurity gas when least wanted, or eating up the hydrogen fuel gas from the reactor.

Acting on the reasonable assumption that after the materials scientists have done their best the plasma will still contain a certain amount of impurity, plasma physicists have designed 'divertors' to clean out the impurities. The principle is to draw off the outer

It is able to predict faithfully accurate data on second virial coefficients, thermal conductivity, thermal diffusion and collision cross sections.

With what must be almost the last word on the Ar-Ar potential having been said, it is possible for experimentalists to extend the scope of the problem and look at more complicated partners, particularly exploiting the properties of lasers which can be used to simplify the spectra of diatomic species.

An example of this is a study of the electronic spectrum of NaNe by Ahmad-Bitar *et al.* (*Phys. Rev. Lett.* **39**, 1957; 1977). The molecule which is only held together by weak Van der Waals forces was created by using the cooling produced on expansion in a supersonic jet. The beam of molecules issuing from the nozzle was crossed at right angles with a laser beam working in a single mode. This procedure avoids Doppler broadening in the resonance fluorescence spectrum which can be observed. Photons from the laser are absorbed and then re-emitted from the molecule without any redistribution of energy in the excited electronic state.

In NaNe, the depth of the potential well was shown to be about 8 cm^{-1} compared with approximately 100 cm^{-1} in the case of Ar_2 , making NaNe the most weakly bound molecule yet studied by high resolution spectroscopy.

Graham Richards is a Lecturer in the Physical Chemistry Laboratory and a Fellow of Brasenose College, Oxford.

skin of the plasma, known as the scrape-off layer, into side chambers through electromagnetic canals. There the impurities are separated from the hydrogen fuel which is eventually returned to the main chamber. The scrape-off layer intercepts the heavy impurities on their way from the walls, and provided it can be skimmed off fast enough by the divertors, the central plasma can be kept clean. The pros and cons of the alternative divertor designs were debated at length last week by the proud owners of tokamaks from Japan, Germany, France, USA and Britain. Clearly these divertors do work and we were shown that they can reduce the concentration of impurities in the main plasma by almost an order of magnitude.

M. W. Thompson is Professor in the School of Mathematical and Physical Sciences, University of Sussex.

There is still a long way to go before fusion can be counted amongst the alternative energy sources. The figure of merit, which is the algebraic product of plasma density, temperature and containment time is still two orders of magnitude from the critical value. But the many problems of developing these reactors are being tackled in the right order and after the conference last week some more of them are known to be soluble. □

Growth factors for tumours

from Robert Shields

A VARIETY of cells transformed by murine or feline sarcoma viruses lose the ability to bind epidermal growth factor (EGF) to cell surface receptors. However, cells transformed by DNA viruses, avian sarcoma viruses or infected with non-transforming RNA viruses show normal binding of EGF (Todaro *et al.* *Nature* **264**, 26; 1976). Two types of explanation can be advanced for these observations. The more prosaic (and more likely) is that the sarcoma gene product directly or indirectly interferes with the amount or function of EGF receptors. Indeed it has been shown recently that mutant cells defective in the glycosylation of cell surface proteins have diminished amounts of EGF receptors on their surfaces and it seems quite possible that transformation by certain sarcoma viruses might have a similar effect (Pratt & Pastan *Nature* **272**, 68; 1978).

The more intriguing proposal put forward by Todaro and his coworkers (for which they produce no evidence) is that virally transformed cells are producing EGF or an EGF-like substance that competes with authentic EGF in the assay for the receptors. Since EGF is a growth factor for a number of cells this could mean that the transformed cells producing EGF would be permanently stimulated to grow. This could account for the lowered requirement for exogenous growth factors often shown by transformed cells. There are precedents for cells producing material that interacts with receptors on their own surface, as lines of neuroblastoma cells both produce and respond to nerve growth factor (NGF) (Bradshaw & Young *Biochem. Pharm.* **25**, 1445; 1976).

In a recent paper (*Nature* **272**, 356; 1978) De Larco and Todaro produce more direct evidence for the synthesis

Robert Shields is at the Imperial Cancer Research Fund Laboratories, London.

Aromatic polyamides

from Paul Calvert

ONE of the early heroes of polymer science is W. H. Carothers who did much of the original synthesis and development of aliphatic polyamides, nylons, and made a great deal of money for his employers, E. I. Du Pont, in the process. Nowadays synthetic fibres make losses rather than profits and a polymer chemist with a new textile fibre would probably be told to go away and lose it. In this light it is encouraging to read a series of papers by P. W. Morgan and his coworkers, also of Du Pont, on aromatic polyamides, which remind us that industrial research can still be both useful and interesting.

In a paper first presented when he was given the Witco Award by the American Chemical Society in 1976, Morgan reviews the synthesis and properties of aromatic and other rigid polyamides (*Macromolecules*, **10**, 1381; 1977). These polymers are of interest because they are resistant to high temperatures, they dissolve in only a few very polar solvents such as concentrated sulphuric acid and amide solvents, they often form liquid crystalline solutions, and if spun into fibres from those solutions they form fibres with very high modulus and strength.

The succeeding three papers in the same issue of *Macromolecules* go on to describe the properties of these liquid crystalline solutions in more detail. The viscosities of solutions of poly(1, 4-benzamide) increase as a function of polymer concentration, then abruptly start to decrease at a

critical value corresponding to the onset of opalescence in the solution. At higher concentrations the liquids consist of a mixture of isotropic solution and more concentrated nematic liquid crystalline solution. The polymer is partly fractionated with more high molecular weight chains in the anisotropic phase. Fibres spun from liquid crystalline solvents and heat treated reach strengths of 4×10^8 psi and moduli of 2×10^7 psi. NMR and infrared studies of poly(1, 4-benzamide) in dimethylacetamide-lithium chloride mixtures suggest that the polymer molecules exist associated to the chloride ions to form a negatively charged chain which pairs with a dimethylacetamide-Li⁺ complex, giving a neutral chain which dissolves. One would like to know a lot more about these associations, particularly because of the analogy with protein solutions. On standing or in a magnetic field the anisotropic phase becomes uniformly oriented and transparent. Under these circumstances one can see ellipsoidal droplets of the isotropic phase, distorted from spherical by the surrounding anisotropic liquid.

These materials are in use as high temperature and reinforcing fibres, but relatively little has been published in the past apparently because of industrial secrecy. These papers should stir up considerable academic interest in these materials.

Paul Calvert is in the School of Molecular Sciences, University of Sussex.

of growth factors by tumour cells. They report that a line of human fibrosarcoma cells produces a growth factor similar to MSA, a growth factor which was first described as the product of a rat liver cell line grown in serum-free medium (Dulak & Temin *J. cell. Physiol.* **81**, 153; 1973). The molecule is a member of the family of closely related hormones, the somatomedins (see *News and Views* **267**, 308; 1977; **270**, 665; 1977); like the somatomedins MSA has non-suppressible insulin-like activity (Dulak & Shing *J. cell. Physiol.* **90**, 127; 1977), a specific binding protein (De Larco & Todaro *op. cit.*) and is a growth factor for cultured cells (Rechler *et al.* *Eur. J. Biochem.* **82**, 5; 1978). Exactly which member of the human somatomedin family is produced by these fibrosarcoma cells will have to await more sophisticated

analysis.

This is not the first report of MSA-like material synthesised by tumours. Patients with various non-islet cell tumours and suffering from hypoglycaemia have elevated serum levels of MSA-like material whose non-suppressible insulin-like activity accounts for the hypoglycaemia (Megyesi *et al.* *J. clin. Endocr. Metab.* **38**, 931; 1974) and the tumours contain particularly high concentrations of the molecule suggesting that it is being produced by the tumour itself (Hyodo *et al.* *J. clin. Endocr. Metab.* **44**, 1175; 1977). The normal site of synthesis of somatomedins is not known although the liver seems a likely candidate (Van Wyk *et al.* in *Advances in Metabolic Disorders* (eds Luft & Hall) **8**, 127 (Academic Press, New York, 1975)). However, such a diverse collection of tumours seem to make somatomedins that some at least

must be making it ectopically (Hyodo *et al. op. cit.*; *News and Views* **269**, 752; 1977). This ectopic elaboration of a growth factor may be more than a clinical curiosity; ectopic protein synthesis may have a role in the growth and development of a number of tumours and help explain the phenomenon of ectopic hormone synthesis itself.

Tumour cells probably arise through a series of changes which give them the competitive edge over their normal counterparts. One way this can be done is to decrease their dependence on serum which normally supplies growth factors and in fact decreased serum dependence often accompanies cell transformation. Another way is to increase the supply of growth factors to the tumour. This can be achieved by their synthesising their own growth factors, increasing the range of growth factor receptors expressed by the cell to take advantage of local conditions or by inducing tumour vascularisation to improve growth factor supply. All these objectives may be achieved through ectopic protein synthesis.

Besides the examples of MSA-secreting tumours mentioned above, a variety of other cells have been reported to manufacture growth factors. For instance, a line of BHK cells transformed by SV40 virus produces a factor that stimulates growth and movement of 3T3 cells (Bürk *Expl Cell Res.* **101**, 293; 1976). This factor appears to be distinct from MSA as its molecular weight (18,000) is higher and unlike MSA it acts synergistically with insulin (Bourne & Rozengurt *Proc. natn. Acad. Sci. U.S.A.* **73**, 4555; 1976). A number of different laboratories have reported that growth factors are found in medium conditioned by L929 cells, these include nerve growth factor (NGF) (Oger *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 1559; 1977), colony stimulating factor (CSF) which is identical to macrophage growth factor (MGF) (Stanley *et al. J. exp. Med.* **143**, 631; 1976) as well as other unidentified growth factors (Shodell *Proc. natn. Acad. Sci. U.S.A.* **60**, 1455; 1972). Since the normal site of synthesis of these factors is uncertain it may be premature to call them ectopically synthesised but nevertheless these results show that a variety of tumour cells can make growth factors. Whether these cells can respond to growth factors they themselves produce is not clear, but since some of them 'condition' tissue culture medium (Nissley *et al. Cell* **11**, 441; 1977), this seems likely.

There is as yet only scanty evidence that tumour cells may synthesise new hormone receptors to take advantage of the prevailing conditions (Cikes *Eur. J. Cancer* **14**, 211; 1978). However,

some data which may be relevant come from the study of the distribution of NGF receptors in a number of tissues. The main site of NGF synthesis in the mouse seems to be the submandibular gland (Berger & Shooter *Proc. natn. Acad. Sci. U.S.A.* **74**, 3647; 1977). However, if the gland is removed, serum levels of NGF soon return to normal (Murphy *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 2330; 1977). The probable explanation for this is that NGF is synthesised systemically by a variety of tissues, an idea bolstered by the observation that an enormous variety of normal and tumour cells produce NGF in culture (Bradshaw & Young *op. cit.*). Whereas most cells seem to possess receptors for growth factors such as EGF, MSA and FGF (fibroblast growth factor) only very few cells related to neural tissue have receptors for NGF (Fabricant *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 565; 1977). Thus a tumour cell might gain competitive advantage not by making NGF but by the ectopic synthesis of NGF receptors; it would then be in a position to benefit from systemic NGF production. There is even some evidence to support this notion. Malignant melanoma cells are unusual in that they have abundant NGF receptors and since pigment cells are descended from the neural crest, melanoma cells may have dedifferentiated to re-express embryonic NGF receptors ectopically (Fabricant *et al. op. cit.*). It would be interesting to know if tumour cells which produce ectopic hormones (such as ACTH) also have receptors for these same hormones. If so, then it is conceivable that ectopic production of a variety of peptide hormones could autostimulate the producing cell in the way proposed by Todaro *et al.*

Unless adequately provided with a blood supply tumour growth is restricted by the supply of nutrients and growth factors as well as the accumulation of waste products. However, a number of tumours as well as cultured tumour cells secrete a factor or group of factors termed tumour angiogenesis factor (TAF) which can induce vascularisation of tumours (Folkman *In Cancer* (ed. Becker, F.) **3**, 355 (Plenum, New York, 1976)). Normal cells do not make TAF which seems to be an ectopically produced growth factor for endothelial cells (Klagsbrun *et al. Cancer Res.* **36**, 110; 1976; Birdwell *et al. Nature* **268**, 528; 1977). Since ectopic proteins are normal products made in unusual amounts in unusual places, it seems likely that TAF's usual role is in vascularisation and endothelial growth under normal circumstances.

The results discussed here show that a plausible explanation for ectopic

protein production is that cells which synthesise ectopic growth factors, make new hormone receptors or induce fresh blood supply through ectopic TAF synthesis may gain a selective advantage over their normal counterparts. Ectopic protein synthesis could then play a vital part in the growth and development of many different kinds of tumour. □

A liquid permanent magnet confirmed

from P. V. E. McClintock

It seems that superfluid $^3\text{He-A}$ really is a liquid ferromagnet, and thus possesses a permanent magnetisation of its own even in the absence of any externally applied magnetic fields. The prediction by A. J. Leggett of Sussex University that this might be the case, published in *Nature* (**270**, 585; 1977) only a few months ago, has apparently now been confirmed by an ingenious experiment carried out by D. N. Poulson and J. C. Wheatley of the University of California at San Diego and reported in a recent issue of *Physical Review Letters* (**40**, 557; 1978).

Leggett's prediction arose naturally from the 'giant diatomic molecule' model of superfluid ^3He in which the liquid is regarded as being composed of pairs of atoms each rotating about the other. In all diatomic molecules there is some tendency for a little of the orbital motion of the atoms to get transferred to motion of electrons around their respective nuclei, constituting circulating currents which thus endow the molecule with a weak magnetic movement. In a conventional diatomic gas, the molecules are randomly orientated with respect to each other, so that the net magnetisation arising from this effect averages to zero. In superfluid $^3\text{He-A}$, on the other hand, the 'molecules' are aligned in a highly ordered state, with the orbital angular momentum of each pair of atoms tending to point in the same direction, described by a vector $\mathbf{l}$. Leggett pointed out that any magnetic moments carried by the individual pairs would thereby reinforce each other, rather than cancelling out, and he predicted that there should consequently be a net magnetisation of the liquid as a whole directed along the $\mathbf{l}$ vector.

Designing an experiment to look for this phenomenon was far from easy. Quite apart from the intrinsic diffi-

P. V. E. McClintock is a member of the Department of Physics at the University of Lancaster.

culty of working at the very low temperatures (around 2 mK) required to create the A-phase, the magnetic field being sought was expected to be exceedingly feeble: around 1/50 of the Earth's field: so that an indirect approach was required. The principle of the experiment rested on a very careful examination of how the $\mathbf{l}$ vector responded to an externally applied magnetic field.

The orientation taken up by $\mathbf{l}$ is determined by the combination of a wide range of influences; in particular, $\mathbf{l}$ tends to lie perpendicular to magnetic fields and parallel (or antiparallel) to the direction of any heat flow in the liquid. It also bends in such a way as to meet the boundaries of the container, and any other solid surfaces, at right angles. It should be emphasised that none of these effects influences the sign of $\mathbf{l}$, but only its direction. Thus, in the absence of Leggett's ferromagnetism, an exact reversal in the direction of an externally applied magnetic field from $\mathbf{H}_{\text{ext}}$ to $-\mathbf{H}_{\text{ext}}$ would be expected to have no effect whatsoever on the orientation of $\mathbf{l}$. If, however, Leggett's prediction was correct so that, in addition to $\mathbf{H}_{\text{ext}}$, a weak magnetisation also existed parallel to $\mathbf{l}$, reversal of $\mathbf{H}_{\text{ext}}$ would not, in fact, produce a precise reversal of the net magnetic field within the liquid. Consequently, the direction of $\mathbf{l}$ would be expected to change slightly towards a position perpendicular to the new direction of the net field. Paulson and Wheatley's experiment consisted, in effect, of looking for tiny shifts in the direction of $\mathbf{l}$ following a reversal of the external magnetic field.

Fortunately, a very sensitive technique was already available for detecting changes in the orientation of $\mathbf{l}$: the attenuation of zero sound (a collisionless form of sound wave peculiar to liquid ^3He) is strongly affected by the angle between $\mathbf{l}$ and the direction of propagation of the sound wave. Thus, with a fixed transmitter and detector, and using an input signal of constant amplitude, the strength of the received signal gives immediate information about the orientation of $\mathbf{l}$ throughout the intervening liquid.

The magnitude of $\mathbf{H}_{\text{ext}}$ was chosen carefully so as to be large enough to be able to influence the direction of $\mathbf{l}$, but not so large as totally to mask the anticipated magnetisation of the liquid. The direction of $\mathbf{H}_{\text{ext}}$ was first adjusted so that the received signal was of the same amplitude as when $\mathbf{H}_{\text{ext}}=0$, indicating that on average $\mathbf{H}_{\text{ext}}$ was then perpendicular to $\mathbf{l}$ throughout the liquid. The direction of $\mathbf{H}_{\text{ext}}$ was then reversed: and a small change in the received signal amplitude was indeed seen, indicating a shift in the direction of $\mathbf{l}$ and, hence, implying that $\mathbf{H}_{\text{ext}}$

was not the only magnetic field acting within the liquid. The authors have done their best to eliminate the possibility that this might have been a spurious effect arising from stray magnetic fields external to the ^3He . They argue, with some justification, that the magnitude of any such stray fields would be temperature independent, whereas the observed effect seemed to vary with temperature in very much the manner expected for the anticipated ferromagnetism of the liquid itself. On being repeated several times the experiment was found to give reasonably reproducible results. The rather non-ideal texture taken up by $\mathbf{l}$ inside the experimental cell made it difficult to quantify the effect, but it was nevertheless possible to establish that the magnetism was approximately of the expected magnitude.

Although further work in a more carefully controlled geometry is clearly needed to sort out all the quantitative details, it is evident that Leggett's prediction has been vindicated and that $^3\text{He-A}$ thus becomes the first ferromagnetic liquid to have been discovered. $\square$

Search for superheavies

from Daphne F. Jackson

At a recent international symposium on superheavy elements* the experimental situation was best summarised by G. Herrmann (University of Mainz) who concluded from his review of attempts to produce superheavy elements in heavy ion reactions that there is no evidence for formation of superheavy nuclei which survive—yet!

Superheavy elements are expected to be found at islands of stability beyond the present known elements in the Periodic Table. There is now reasonable agreement that the elements with atomic nuclei having neutron number $N=184$ and proton number $Z=110$ or 114 are most likely to be observed although the nucleus $N=228$, $Z=126$ is not excluded. These nuclei should be spherical; the prospects for observing deformed superheavy nuclei are not good as such nuclei are predicted to be unstable against β -decay.

Superheavy nuclei are not expected to be stable. This means that the prospects for finding the corresponding elements in nature or of detecting their formation in a nuclear reaction depend on the half-lives for α -decay and fission and the subtle competition between

these two processes. The calculation of half-lives for α -decay, β -decay and fission were reviewed by S. G. Nilsson (Lund) and A. Sobiczewski (INR, Poland). Nilsson pointed out that the calculation of α -decay rates depends essentially on predictions of the nuclear masses and questioned the reliability of mass formulae. He argued that studies of the deformed actinide nuclei give a better test of mass calculations than in the closed shell region around lead where predictions of both masses and single-particle levels are incorrect.

Calculations of the position and ordering of single-particle levels and consequent prediction of N and Z for closed shells were discussed by R. Chasman (Argonne National Laboratory), F. Petrovich (Florida State University), M. A. K. Lohdi (Texas Tech University) and P. Quentin (Los Alamos Scientific Laboratory). Chasman and Lohdi emphasised the role of the non-locality of the single-particle potential while Petrovich stressed the significance of the spin-orbit interaction in determining shell closure. Most speakers agreed on the importance of correctly reproducing the single-particle behaviour and shell corrections in the actinides, although this view was challenged with the comment that the potentials preferred by Petrovich yield incorrect results for overall nuclear sizes in the lead region.

Important contributions on the behaviour of fission barriers and fission half-lives came from M. Mustafa (Lawrence Livermore Laboratory) and F. B. Malik (Indiana University). Mustafa considered the dependence of the potential energy surface on the deformation parameters and pairing gaps, which in turn depend on the angular momentum and excitation energy of the system. He showed that the fission barrier disappears for angular momentum greater than 32 units and excitation energy greater than 50 MeV, and hence if the superheavy nucleus is formed in this condition, for example in a heavy ion collision, it will break up immediately. Malik presented results derived using a variable density for the overlapping nuclei and showed that this model predicts the existence of an external barrier of a few tens of MeV between the saddle point and the scission point. In this theory the fission lifetimes of all superheavy nuclei around $Z=114$ are considerably shorter than 1 s.

Methods for calculating half-lives for α -decay were reviewed by D. F. Jackson (University of Surrey). She showed that it is possible to eliminate uncertainties due to channel radii and ambiguities in α -nucleus potentials but stressed the importance of antisymmetrisation. Nevertheless, the new results presented for superheavy nuclei

* Held at Texas Technological University, March 9-11, 1978.

agreed to within an order of magnitude with the semi-empirical formula of Viola and Seaborg. Much more serious discrepancies were reported by S. Wald (Texas Tech) between exact calculations of fission half-lives and the conventional approximate calculations. The exact calculations yield shorter half-lives by several orders of magnitude.

Herrmann reviewed experimental work on collisions of uranium ions on uranium and xenon on uranium, and gave upper limits on cross sections for formation of superheavy nuclei. K. Hulet (Lawrence Livermore Laboratory) described a systematic study of the interaction of ^{48}Ca ions on ^{248}Cm and again gave very low cross section limits for the formation of superheavy elements. Both J. M. Nitsche (Lawrence Berkeley Laboratory) and J. R. Otto and colleagues also from the Lawrence Berkeley Laboratory challenged the widely held view that $^{48}\text{Ca} + ^{248}\text{Cm}$ and $^{136}\text{Xe} + ^{248}\text{U}$ reactions are the most suitable for reaching new islands of stability. It seems likely that present estimates for complete fusion cross sections must be re-evaluated.

Searches for superheavy nuclei in nature and in accelerator laboratories frequently focus on detection of spontaneous fission, particularly as it has been predicted that the average number of neutrons per fission is higher than in known nuclei. Meteorites and radiogenic lead have been studied in this way. However, D. C. Hoffman (Los Alamos) argued, from the basis of work on californium and fermium isotopes, that the neutron emission in spontaneous fission of superheavy elements could be low. She also stressed the importance of careful chemical separation.

Evidence was presented for the observation of long range α -particles in such samples as Madagascan monazite and biotite, meteorites, brine from hot springs, nuclear emulsion plates, and CERN targets. The significance of these results was discussed with vigour, but without conclusion.

In his opening review talk O. L. Keller (Oak Ridge National Laboratory) noted that the atomic properties of superheavy elements are well understood compared with the nuclear situation. It is possible to give reliable predictions of intensities and energies of X rays emitted by superheavy atoms; further, if a million such atoms could be produced it would be possible to use synchrotron radiation to excite X-ray emission in an extension of Moseley's historic experiments. Further discussions of atomic effects seen in heavy ion collision were by W. E.

Meyerhof (Stanford University) and W. Wölfi (University of Zurich).

The conclusion of this conference must be that convincing experimental observation of superheavy elements is probably not imminent. The search is, however, leading to vital new understanding of nuclear and atomic physics and quantum electrodynamics. $\square$

Dynamical diseases

from Robert M. May

THERE have been several accounts in these columns of the extraordinary variety of dynamical behaviour that can arise from nonlinearities in simple deterministic systems (see *News and Views* 256, 165; 1975; 271, 305; 1978 and *May Nature* 261, 459; 1976). Such a range of behaviour, from stable points, through stable cycles, to apparently random fluctuations, can be exhibited by a first-order differential equation with time lags, or by systems of three or more coupled, ordinary first-order differential equations.

Recently, it has been suggested that many human illnesses may be associated with physiological control processes that are in one regime of dynamical behaviour for normal, healthy people, and in another regime for sick people. Typically, the physiological variable may be constant in the normal state, and oscillatory or 'chaotically' varying in the pathological state. Mackey and Glass (*Science* 197, 287; 1977) have christened such afflictions 'dynamical diseases'.

These authors give a simple first-order differential-delay equation that may plausibly describe the concentration of CO_2 in the blood. There is a steady production rate of arterial CO_2 , and a loss rate that depends on the ventilation (assumed to be a sigmoidal function of the arterial CO_2 concentration) and on the time lag, τ , between oxygenation of blood in the lungs and stimulation of chemoreceptors in the brainstem. In this model, there is a steady solution (corresponding to constant CO_2 concentration), which gives way to increasingly severe oscillations as either τ or the steepness of the CO_2 response curve increases. Mackey and Glass suggest this bifurcation into oscillatory behaviour may correspond to the Cheyne-Stokes phenomenon, characterised in humans and in dogs by regular waxing and waning of breathing amplitude.

Mackey and Glass estimate that oscillatory behaviour will arise once a certain combination of parameters

(characterised by S_0) exceeds a critical value around 7.4 litres per min per mm Hg. In normal people, S_0 lies in the range 2–6, and it is higher in Cheyne-Stokes patients; a similarly crude estimate of the period of Cheyne-Stokes breathing gives a theoretical figure of one-half to one-third the observed period. In a later review (Glass & Mackey, *Annls N.Y. Acad. Sci.*, in the press), the authors survey several other respiratory disorders that may have a similar dynamical character.

A second class of diseases discussed in Mackey and Glass's *Science* paper involves the production of the formed elements of the blood (white cells, red cells, platelets) in the bone marrow. There are a variety of well-documented pathologies characterised by periodic or chaotic fluctuations in the populations of these formed elements, as well as pathologies in which densities remain approximately constant at abnormal levels. Mackey and Glass give two alternative differential equations as plausible descriptions of the processes regulating the population density of circulating blood cells; both equations include a time lag, τ , between production in the bone marrow and release into the bloodstream. As τ lengthens, both these equations manifest bifurcation from a steady state into periodic oscillations, and one of them exhibits further bifurcations into a regime of chaotic fluctuations. Again, necessarily crude estimates of the parameters suggest that these theoretical insights are not qualitatively discrepant with the clinical facts.

In a considerably more abstract paper, Cronin (*Bull. math. Biol.* 39, 187; 1977) has considered the general sort of coupled, ordinary first-order differential equations that may describe the biochemical processes underlying periodic catatonic schizophrenia. She shows that under reasonable assumptions one can end up with "the approximately periodic changes in the concentration levels of the thyroid which are associated with approximately periodic changes in the symptoms in periodic catatonic schizophrenia". More generally, the analysis suggests that the emotional symptoms and the biochemical levels are likely to vary in a random or unpredictable manner (once the pathological region of parameter space has been entered). Such unpredictability or random variability in the symptoms of schizophrenia are known to occur.

As yet, these mathematical studies bear a metaphorical relation to the real diseases. But they provide an inter-

esting way of looking at things, and could lead to more substantial progress before long. As Mackey and Glass observe, if we could really tie a disease to a specific description of some underlying physiological control process, then the search for therapies could be guided by the aim of manipulating control processes back into the normal range.

Ultraviolet lasers for purer silicon

from Andrew Holmes-Siedle

THE range of wavelengths at which laser action can be obtained has been widening as research on new laser materials has continued. That range now extends from the far infrared to the far ultraviolet. Engineers and spectroscopists have been making important advances by utilising the coherence, high intensity and low divergence of the laser in the visible and infrared. Chemists, in general, have had to wait until recently to obtain photochemical effects because few bonds are excited at wavelengths longer than the violet (410 nm, corresponding to a photon energy of ~ 3 eV). A few rare-gas lasers have been found which operate at very short wavelengths (xenon, krypton and argon at 173, 146 and 126 nm respectively). Since 1975, however, much greater interest has been aroused amongst photochemists by the appearance of a new class of laser—the rare-gas monohalide or excimer type. In these the emitting species is an unusual compound, formed only after excitation of a rare-gas atom, which can then combine with fluorine atoms. For example, the compound of argon would be written as ArF*, an excimer. The fact that this compound cannot exist in the ground state is important in the recycling of atoms to regenerate the excited state after emission. The net result of this unusual form of excitation is that the conversion of a pulsed electrical discharge to an intense, monoenergetic light beam of high photon energy is more efficient than obtained in most lasers. The wavelengths produced are rather conveniently spread about the near ultraviolet as follows

ArF* 193 nm

KrF* 248 nm
XeF* 350 nm

It is an indication of the intense interest that these sources arouse in chemists that these lasers are already available in a neat commercial package, yielding up to 100 mJ per pulse, and capable of many pulses without refill.

Perhaps the most obvious advantages of such a beam are that one may be able to produce chemical reactions very selectively because of the concentration of the beam energy into one wavelength. Up to now, flashlamps have had broad spectra and, even though one could filter out all but one band, one would lose a very large proportion of the available energy in the process. Two workers at the Los Alamos Scientific Laboratory have found a reaction which serves as a very good candidate for such an experiment. In *Applied Phys. Letters* (32, 46; 1978), John Clark and Robert Anderson describe the selective removal of impurities from silicon tetrahydride, SiH₄, a gas commonly known as silane. This compound is being used increasingly as an intermediate in fabricating silicon films for semiconductor devices, in which extreme purity is mandatory.

The separation depends on two features: silane does not strongly absorb light in the region of 200 nm wavelength, while the main impurities do so, and the photolysis products of the contaminants precipitate out from the gas. Thus, in principle, one can irradiate the mixture until all the impurities are precipitated. Such a form of purification is quite attractive and rare. Most separation methods rely on 'fractionation', in which the purification step (distillation for example) removes only a fraction—albeit a large one—of the impurities per step and hence several cycles may be necessary, to the detriment of yield of the product. In the present case, we can, in theory, remove the contaminants to the last molecule.

In order to design such a separation, one must have good measurements for the energy absorption and reaction rates of the compounds concerned. To measure these the authors had to overcome some practical difficulties such as the reaction of the gases with the walls of the spectrometric chamber, in order to obtain optical absorption coefficients but found that, in the range 190 to 200 nm, the absorption cross sections for the important impurities in silane, PH₃, AsH₃ and B₂H₆, were about 10,000 times higher than for silane, and that the photolysis products were indeed removed from the gas stream.

So far, only fairly simple separations have been carried out. The authors ir-

radiated mixtures with ratios 10:1 and 100:1 of one hydride along with SiH₄, using an ArF* laser, of wavelength 193 nm. The changes in concentration were measured by sampling the gas stream with a gas chromatograph. After many laser pulses, the concentration could be reduced by 99% for AsH₃ and 40% for the other impurities. There is no reason to believe that one could go much further by longer or more intense irradiation. A small amount of silane was also lost but the proportion was probably not significant. The authors imply that this type of process promises unusual efficiency in energy use when compared with that used in competing purification methods.

The fact that silane is of growing use in making large-area thin-film electronic layers makes this finding one of some possible future technological significance. Only recently (Wilson *et al. Nature* 272, 152; 1978), the fabrication of 4.8% efficient thin-film solar cells by the deposition of silicon from silane gas was reported, and p-n junctions in amorphous silicon have also been made from the same source (see *News and Views* 260, 667; 1976). The greater control one can obtain over the impurities in the semiconductor layer, the greater scope one has for design of new device principles. It is thus encouraging to see that even at this early stage of development, ultraviolet lasers have made a contribution to a key area of purification technology. □



A hundred years ago

COMING TOTAL SOLAR ECLIPSE

There is no doubt whatsoever that the eclipse which will sweep over the United States next July will be observed as no eclipse has been observed before. The wealth of men, the wealth of instruments, and the wealth of skill in all matters astronomical, already accumulated there, makes us Old Country people almost gasp when we try to picture what the golden age will be like there, when already they are so far ahead of us in so many particulars.

From *Nature* 17, 18 April, 481; 1878.

Andrew Holmes-Siedle is a consultant to the Physics Department of the Fulmer Research Institute.

review article

The cosmic-ray isotopes

Peter Meyer*

Cosmic rays provide information on the composition and structure of the Galaxy. Recent measurements suggest that the isotopic composition at the cosmic-ray sources is similar to that of the Solar System and that cosmic-ray acceleration follows nucleosynthesis after a considerable delay. Cosmic rays may therefore be a sample of the present interstellar medium rather than fresh supernova ejecta. Observations of the abundance of radioactive isotopes lead to a containment time of cosmic rays in the Galaxy an order of magnitude larger than had been suspected.

SPECTROSCOPIC analysis has proved to be a most important quantitative tool in all branches of the astronomical sciences. This applies not only to the electromagnetic radiations, but also to the study of the astrophysical aspects of cosmic-ray research. The most significant progress in understanding the origin and propagation in the Galaxy of the energetic cosmic-ray particles has emerged from investigating their energy spectrum, and the spectrum of the species of which they are composed. Over the past 15 years it has been recognised that the cosmic rays are composed of nuclei that encompass the entire range of the known chemical elements—from hydrogen to the actinides—and that they include electrons and positrons. Most important, measurements of the elemental abundance distribution and their similarity with the Solar System abundances have provided the evidence that nucleosynthesis processes in stellar interiors must have been responsible for producing the observed distribution. Although the mechanisms that accelerate the cosmic-ray particles to their enormous energies are not yet understood, there is no doubt that the place of their birth lies in the interiors of stars, the sites where, as Hoyle was first to suggest, all elements, with the exception of the very lightest ones, originate.

In the metabolism of the galaxies, stellar and interstellar materials become eventually quite thoroughly mixed, and it is not immediately obvious at what stage the unique sample of material that forms the cosmic rays is accelerated to its high energies. Because cosmic rays are the most energetic of all astrophysical radiations, it has seemed natural to suggest that they originate in the most energetic processes known in the Galaxy—supernovae. Much of the modern work in cosmic-ray research is aimed at experimentally establishing the connection between the supernova process and cosmic rays and, in doing so, to test models of supernovae and nucleosynthesis mechanisms.

The cosmic-ray particles observed in the vicinity of Earth are not an unadulterated sample of material emerging from their sources. About half of the cosmic-ray nuclei that are heavier than helium suffered collisions with the tenuous interstellar material before reaching the observer. The source nuclei are therefore depleted, spallation products are created, and the observed charge spectrum is considerably different from the abundance distribution at the sources. Figure 1 (from ref. 1) shows the cosmic-ray charge spectrum as observed near Earth and shows, for comparison, the Solar System abundance distribution of the elements, which can be considered the best approximation to a universal abundance². Clearly, the elements Li, Be, B which, for reasons that are well

understood have a low universal abundance, are abundant in the cosmic rays. They are spallation products, originating mainly from the break-up of C and O. Similarly, most of the elements between $Z = 17$ and 25 are secondaries from the break-up of Fe nuclei emitted by the sources. The information contained in the elemental abundance distribution is therefore twofold. The spallation products provide insight into the questions of cosmic-ray propagation and containment, and hence into the nature of the interstellar medium. The abundance of source nuclei, after proper correction for the effects of propagation, are used for diagnosing the nature of the sources.

Elucidation of the elemental distribution of the cosmic rays has provided considerable insight into their origin and propagation, but important advances are possible through the determination of the 'fine structure' of the abundance distributions—the isotopic abundances. Both nucleosynthesis and spallation involve nuclear processes rather than atomic processes and the determination of abundances of individual nucleides, that is isotopes, is therefore crucial if one wishes to pursue the problems of origin and propagation of cosmic rays. Hence, isotopic abundance studies have become one of the central activities in the field of experimental cosmic-ray research. In the following I discuss some aspects and implications of isotopic abundance measurements.

Cosmic-ray sources

As early as 1934 Baade and Zwicky⁴ suggested that cosmic rays originate in supernova explosions. These suggestions were made on the basis of energetics. The supernova explosion is the most violent process observed in the Galaxy and is, therefore, a candidate for particle acceleration. The total average energy output of supernovae and the rate of supernova occurrence provide an energy input sufficient to maintain the high energy density of cosmic rays in the Galaxy. Since the time of these early suggestions much progress has been made in understanding the supernova phenomenon. Supernovae are the likely birthplaces of neutron stars and black holes, they seem to be the sites for major stages of nucleosynthesis, and they accelerate electrons to very high energies, as evidenced by the emission of synchrotron radiation from supernova remnants. While there is no direct evidence for the acceleration of nuclei at present an answer may come in the future from γ -ray astronomy through the detection of π^0 γ rays from a young supernova. An enhanced flux of such γ rays would demonstrate the acceleration of nuclei in much the same way that synchrotron radio emission demonstrates the acceleration of electrons.

If the supernovae are indeed the sites of explosive nucleosynthesis, and if the cosmic rays are ejected in the supernova

*Enrico Fermi Institute and Department of Physics, The University of Chicago, Chicago, Illinois 60637.

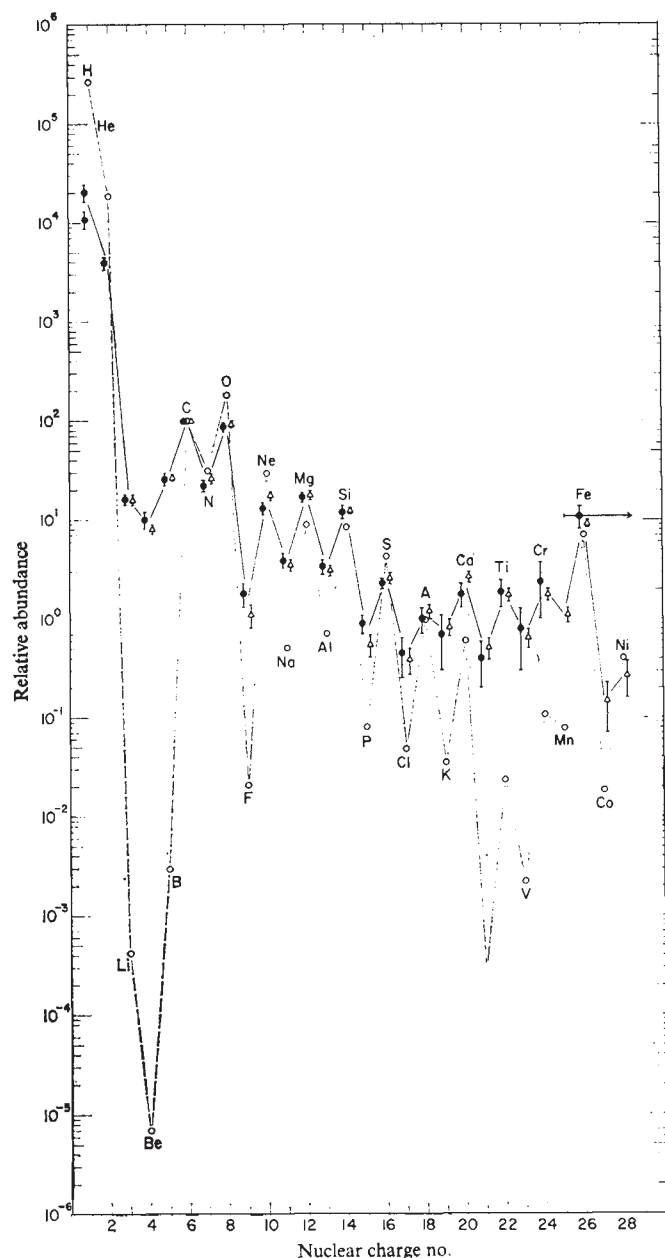


Fig. 1 The elemental abundance distribution of the cosmic radiation and the Solar System abundances, normalised to C. (After Cartwright *et al.*¹ and Cameron²). ●, Measured by University of Chicago cosmic-ray telescope on board the IMP-4 and IMP-5 satellites from 100 to 300 MeV per nucleon; △, Webber *et al.*³ from 250 to 850 MeV per nucleon; ○, Solar System or 'universal' abundances².

process, their composition provides a valuable probe for this process. In particular, measurements of isotopic ratios are important, as such ratios are not likely to be affected by acceleration mechanisms. The crucial parameter that isotopic ratios may reveal is the neutron excess. As thermonuclear burning converts hydrogen to helium, and further fusion processes create nuclei of higher atomic mass, the proton to neutron ratio depends importantly on the density at which the synthesis occurs. At increasing densities, weak interactions become more and more important, and electron capture by nuclei increases the average neutron to proton ratio. The neutron excess is therefore a measure of the maximum density to which the ejected matter has been subjected.

Four nuclear species which are produced copiously by the cosmic-ray sources and whose abundance is not greatly influenced

by spallation processes lend themselves to this kind of study. They are Ne (with stable isotopes at masses 20, 21, 22), Mg (with isotopes 24, 25, 26), Si (with isotopes 28, 29, 30), and S (with isotopes 32 and 34). The existing measurements of the mean neutron excess for these and other elements are shown in Fig. 2 (from ref. 5). Also shown in Fig. 2 is a histogram which is the result of a propagation calculation and gives the mean neutron excess expected to be observed near Earth, if the isotopic abundance of the elements at the sources is the same as that of Solar System material. Within the errors of present experimental techniques, the measurements demonstrate that the three test species for which data exist, Ne, Mg and Si, all exhibit a neutron excess which is compatible with that of Solar System material.

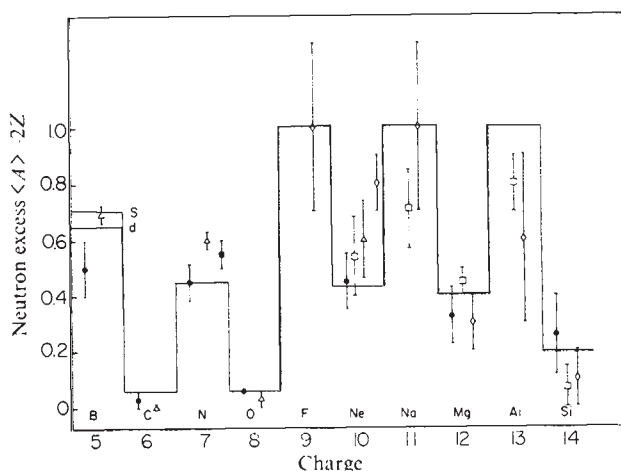
This result leaves two possibilities for the origin of the particles. While it is possible that the nuclei in question are supernova ejecta, their isotopic composition allows the alternative that they are interstellar material accelerated either in conjunction with the supernova process, or by other means not yet understood. Clearly, more accurate observations of the abundance of individual isotopes in this group of nuclei will help to solve this problem.

The most important nuclei for testing the hypothesis of cosmic ray supernova origin are the isotopes of the iron group. Iron itself is among the most abundant of the heavier elements. Since all species heavier than iron and nickel are extremely rare in the cosmic rays, iron is practically free of contamination by spallation of heavier nuclei.

The nucleosynthesis processes that lead to the iron group nuclei have been theoretically investigated in great detail. They are the product of complete silicon burning. Their isotopic composition probes the 'mass cut', that is the boundary between collapsing and ejected material in the explosive supernova process, and it yields information on the temperatures at which the synthesis occurs in the deep regions of the supernova explosion. If the expansion is explosive, and rapid cooling freezes the abundance pattern, this pattern will sensitively exhibit the neutron excess in the high-density regions. This equilibrium process (e process) was originally proposed by Burbidge *et al.*¹¹. The resulting isotopic distribution of Fe as a function of neutron excess has been calculated by Hainebach *et al.*¹² and is shown in Fig. 3 (W. D. Arnett, personal communication).

At an average $Z/A = 0.5$, that is an equal number of neutrons and protons, ^{56}Ni , which subsequently decays into ^{56}Fe by

Fig. 2 The measured average neutron excess of the elements from B through Si in the cosmic rays. The histogram shows the expectation if the isotopic abundance at the cosmic-ray sources is the same as that in the Solar System (after Dwyer and Meyer⁵). ●, Dwyer and Meyer⁵; △, Preszler *et al.*⁶; ◇, Fisher *et al.*⁷; ■, Garcia-Munoz *et al.*⁸; □, Webber *et al.*⁹; —, Shapiro *et al.*¹⁰.



K-capture, is abundantly produced, leading to a predominance of ^{56}Fe . As the neutron excess increases, ^{54}Fe becomes the more abundant isotope, then ^{56}Fe , and finally ^{58}Fe for the largest neutron excess.

Figure 3 demonstrates the motivation for the concentrated effort on the part of experimenters to determine the abundance of the iron isotopes in the cosmic rays. Unfortunately, such experiments are exceedingly difficult and are not yet at a level of sophistication necessary to yield the accuracy required to draw definitive conclusions. However, experiments are being prepared that will permit unambiguous answers to these important questions within a few years.

The Fe, Co and Ni group can provide further information on cosmic ray origin. A different, and less well understood version of the e process creates the isotopes ^{60}Ni , ^{61}Ni , and ^{62}Ni . This process is sensitive to the time scale and character of the explosive expansion. While a measurement of the abundance distribution of the nickel isotopes would be extremely interesting, it requires greatly advanced experimental techniques that are not likely to be available for several years.

A most interesting outcome of observations on cosmic ray isotopes is the determination of the span of time that elapses between nucleosynthesis and acceleration of the particles. This has an immediate bearing on the question whether explosive nucleosynthesis and particle acceleration occur simultaneously, or whether these two phases in the production of cosmic rays are separate. During nucleosynthesis, several nucleides that may only decay by K-capture are created. As long as these particles remain at low energies, their nuclei are surrounded by K-shell electrons, and K-decay can occur. If they acquire high energies after their production (several 100 MeV per nucleon or more) the K electrons are stripped and the energetic nucleus remains stable. Their survival therefore, depends on the time elapsed between production and acceleration. Examples of isotopes that lend themselves to this chronology are ^{57}Co , ^{44}Ti and ^{59}Ni with K-capture decay rates of ≈ 1 yr, ≈ 50 yr, and $\approx 10^5$ yr respectively. Their absence in the cosmic-ray flux would therefore indicate that the time between nucleosynthesis and acceleration would be greater than 1, 50 or 10^5 yr respectively. Good evidence already exists on the abundance of Co, indicating that the timespan between nucleosynthesis and acceleration is longer than 1 yr (refs 14–16).

Cosmic-ray propagation and confinement

As pointed out earlier, a significant fraction of the cosmic rays undergoes nuclear collisions while travelling through the interstellar medium and before reaching the observer. The study of nucleides that emerge after these interactions provides some unique information on the properties of the containment volume and hence, the conditions in interstellar space. It has been known for several years that cosmic-ray particles around 1 GeV per nucleon traverse on the average $5\text{--}6\text{ g cm}^{-2}$ of material before being observed. This evidence was gained through observations of the elemental abundance of those elements that can reliably be assumed to be absent or very rare in the sources, that is, Li, Be, B and most of the elements between $Z \approx 17$ and 25. The former are produced essentially by the spallation of C and O; the latter by the spallation of Fe.

More recently it was discovered that the abundance ratio of the secondary nuclei to their primary progenitors decreases with increasing energy (compare refs 17–21). This phenomenon implies that the particles with higher kinetic energy traverse less interstellar matter than those of lower energy. Several explanations of this effect are possible, the simplest being an energy dependent escape path length and containment time in the Galaxy.

Observations of isotopes add a powerful tool for the study of these topics. The first isotopic ratio that was successfully studied was $^3\text{He}/^4\text{He}$, ^3He being entirely of secondary, and ^4He of primary origin. The approximately 10% admixture of ^3He around 100 to 1,000 MeV per nucleon confirms the value

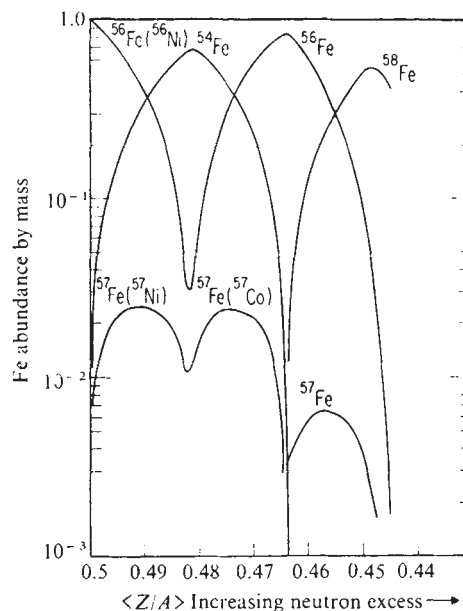


Fig. 3 The isotopic distribution of Fe by mass as a function of neutron excess (W. D. Arnett, personal communication).

for the average leakage path length obtained from the Li, Be and B observations. A measurement of the ^3He fraction at much higher energies has still to be made.

The nitrogen isotopes have, in recent years, been successfully measured over the energy range from a few MeV to over 1 GeV per nucleon. Above a few hundred MeV per nucleon, ^{14}N and ^{15}N were found to have almost the same abundance, a result that is expected if nitrogen is of spallation origin since the cosmic-ray sources would almost exclusively produce ^{14}N . The measurements made it possible to put an upper limit on the primary nitrogen flux of only a few per cent of that of carbon²².

A most striking result was obtained for the nitrogen isotopic abundance in range 3–14 MeV per nucleon. In this energy regime, ^{14}N was found to be considerably more abundant than ^{15}N (ref. 23), confirming a theoretical prediction that these low energy particles most likely are a sample of interstellar gas that becomes ionised and accelerated within interplanetary space²⁴.

Let me finally deal with the cosmic rays as an ingredient of the interstellar medium. To maintain their energy density requires a substantial energy input in the special form of high energy particles. Since this energy density is comparable to that of the galactic magnetic fields and of the turbulent motion in the Galaxy, they must influence the dynamics of the Galaxy. In this context, the determination of the time for which they are contained in the Galaxy is of foremost significance. Radioactive isotopes, serving as clocks, are the most important tool to determine that time. Among the isotopes whose half lives are comparable to the expected galactic containment time are ^{10}Be ($\tau = 1.6 \times 10^6$ yr), ^{26}Al ($\tau = 7.4 \times 10^5$ yr) and ^{36}Cl ($\tau = 3 \times 10^5$ yr). At present, ^{10}Be abundance measurements have been advanced to an accuracy that provides definite conclusions on the cosmic-ray containment time. All Be isotopes are of secondary origin, their relative production yields being determined predominantly by the spallation of C and O. Three Be isotopes occur in this secondary flux: ^7Be , which is stable at cosmic-ray energies since it can only decay by K capture, the stable ^9Be and the radioactive ^{10}Be .

Several investigators have recently presented experimental data on the relative abundances of the Be isotopes^{6,25,26}. As an example of these efforts, I show results from the University of Chicago group that were gained with instruments on the IMP satellites. Figure 4a displays the mass distribution for the

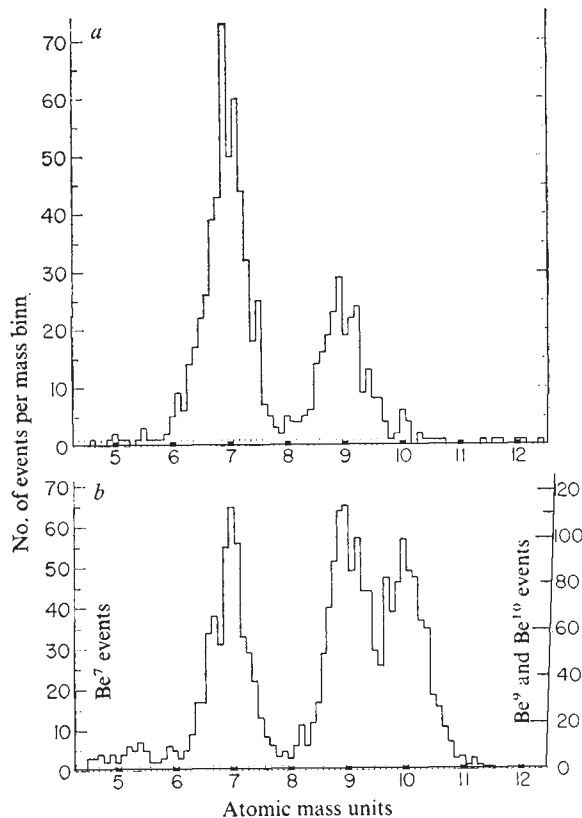


Fig. 4 *a*, Mass distribution of Be observed in the cosmic rays (IMP-7 and IMP-8 1973–76 data). *b*, Mass distribution of Be produced by spallation from an accelerator nitrogen beam¹³.

element Be as observed in the cosmic rays at energies between 30 and 150 MeV per nucleon and shows the striking absence, or near absence, of the isotope ¹⁰Be. The same detector, when exposed to a flux of Be produced by spallation of a nitrogen beam from an accelerator, responds as shown in Fig. 4*b*. A distribution of this general form could be expected in the cosmic rays had all ¹⁰Be survived. The authors reach the conclusion, confirmed by other measurements, that the cosmic-ray containment time is in excess of 10⁷ yr. This has the consequence that the particles must have propagated in a medium of average density of about 0.2 hydrogen atoms cm³, if during that containment time they traversed the well established 5 to 6 g cm⁻² of matter.

Since this value for the interstellar density is considerably below what is generally assumed to be the average density in the galactic disk, one must raise the question of where the particles spend a major fraction of their lives. The answer to this question is by no means clear. An old suggestion, originally proposed to explain the distribution of synchrotron radiation by cosmic-ray electrons has been resurrected. It involves the existence of a low density region around the galactic disk—often referred to as a galactic halo. More recent work has shown that such regions may be produced by gas driven under the cosmic-ray pressure in a manner described by Parker²⁷ (see also ref. 28). Alternatively, the question was raised whether large-scale inhomogeneities in interstellar density exist within the galactic disk that provide volumes filled with high temperature, low density gas in which the cosmic rays spend a major fraction of the time for which they are contained^{29,30}. It is likely that this problem is intimately connected with the observation of a strongly reduced amount of matter that the very energetic cosmic rays traverse before escape and that was mentioned earlier.

On the experimental side, much remains to be done. A most intriguing possibility is a measurement of the ¹⁰Be abundance at highly relativistic energies. At about 10 GeV per nucleon,

for example, time dilatation would lead to a decay time about 10 times that in the rest frame. It is, therefore, possible that a substantial flux of ¹⁰Be exists at those energies. Its determination may yield the actual cosmic-ray containment time rather than an upper limit. The isotopes ²⁶Al and ³⁶Cl have not as yet been isolated to permit their use in cosmic-ray chronology. In both these areas, experiments are in preparation that can be expected to substantially improve our knowledge within a few years.

Progress and problems

Cosmic rays provide us with a unique, observable sample of galactic matter. The determination of their elemental and, more recently, their isotopic composition, will be invaluable in the elucidation of the origin of the elements. Theoretical studies of nucleosynthetic processes have made rapid progress in recent years. Experimental tests for these theories are scarce. The theoretical estimates of relative isotopic abundances are considerably more reliable than those of elemental abundances which, in addition, may be distorted by charge dependent acceleration processes. Accurate determinations of isotopic abundances are therefore of foremost importance.

The first significant conclusions from the work on isotopes are now emerging. Ne, Mg, and Si exhibit an isotopic composition similar to that in the Solar System. The Co abundance points to a considerable time delay between nucleosynthesis and acceleration. These results begin to suggest that explosive nucleosynthesis and cosmic ray acceleration are not as closely related as had been suspected. More accurate determination of the iron isotopes and of the abundances of the elements beyond iron will soon further clarify this question. Should it turn out that the cosmic rays are a sample of interstellar gas, rather than fresh supernova ejecta, the consequences are of similar interest. In that case, the cosmic ray composition will yield meteorite-quality data for the present composition of the interstellar medium, not that of 4.5 × 10⁹ yr ago.

Radioactive clocks are the most important source of information on the containment of cosmic rays in the Galaxy. The recent work on the abundance of the nuclide ¹⁰Be demonstrated that the containment time is more than 10⁷ yr, about a factor 10 longer than previously accepted. This has immediate consequences for models of the structure of the Galaxy, and in particular, the interstellar medium. Any model must provide for the traversal of an average of 5 to 6 g cm⁻² of material at an average density of about 0.2 g cm⁻³ for the 50 to 1,000 MeV nucleon particles.

This brief account of the part that cosmic ray isotope investigations play in astrophysics cannot be comprehensive. Nor is the list of references. Among the important missing topics are the very puzzling isotopic compositions of energetic particles emitted by the sun, and the role that measurements of isotopes play in determining particle propagation and acceleration processes in interplanetary space. The interested reader is referred to two recent summaries of the field (refs 31 and 32).

I thank W. D. Arnett for useful discussions, and NASA (grant NGL-14-001-005) and the US NSF (grant AT 77-04809) for their support.

1. Cartwright, B. G., Garcia-Munoz, M. & Simpson, J. A. *Proc. 12th Intern. Conf. Cosmic Rays* 1, 215 (1971).
2. Cameron, A. G. W. *Space Sci. Rev.* 15, 121 (1973).
3. Webber, W. R., Dample, S. V. & Kish, J. *Astrophys. Space Sci.* 15, 245 (1972).
4. Baade, W. & Zwicky, F. *Proc. natn. Acad. Sci. U.S.A.* 20, 254–259 (1934).
5. Dwyer, R. & Meyer, P. *Astrophys. J. Lett.* 216, 646 (1977).
6. Presler, A. M., Kish, J. C., Lezniak, J. A., Simpson, G. & Webber, W. R. *Proc. Fourteenth Intern. Conf. on Cosmic Rays*.
7. Fisher, A. J., Hagen, F. A., Maehl, R. C., Ormes, J. F. & Arens, J. F. *Astrophys. J.* 205, 938 (1976).
8. Garcia-Munoz, M., Mason, G. & Simpson, J. A. *Symposium on the Isotopic Composition of Solar and Galactic Cosmic Rays* (Durham, New Hampshire, 1974).
9. Webber, W. R., Lezniak, J. A., Kish, J. C. & Simpson, G. A. *Astrophys. J. Lett.* 18, 125 (1977).
10. Shapiro, M. M., Silberberg, R. & Tsao, C. H. *Proc. 14th Intern. Conf. Cosmic Rays* 2, 532 (1975).
11. Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F. *Rev. mod. Phys.* 29, 547 (1957).
12. Haebach, K. L., Clayton, D. D., Arnett, W. D. & Woosley, S. E. *Astrophys. J.* 193, 157 (1974).
13. Garcia-Munoz, M., Mason, G. M. & Simpson, J. A. *Astrophys. J.* 217, 859 (1977).

14. Juliusson, E. & Meyer, P. *Astrophys. J.* **201**, 76 (1975).
15. Meyer, P. & Minagawa, G. *Proc. Fifteenth Intern. Conf. on Cosmic Rays* **1**, 248 (1977).
16. Tueller, J., Love, P., Epstein, J. W., Israel, M. H. & Klarmann, J. *Proc. Fifteenth intern. Conf. on Cosmic Rays* **1**, 254 (1977).
17. Juliusson, E., Meyer, P. & Müller, D. *Phys. Rev. Lett.* **28**, 926 (1972).
18. Smith, L. H., Buffington, A., Smoot, G. F., Alvarez, L. W. & Wahlig, M. A. *Astrophys. J.* **180**, 987 (1973).
19. Balasubrahmanyam, V. K. & Ormes, J. F. *Astrophys. J.* **186**, 109 (1973).
20. Juliusson, E. *Astrophys. J.* **191**, 331 (1974).
21. Caldwell, J. H. *Astrophys. J.* **218**, 269 (1977).
22. Dwyer, R. & Meyer, P. *Phys. Rev. Lett.* **35**, 601 (1975).
23. Mewaldt, R. A., Stone, E. C. & Vogt, R. E. *Proc. Fifteenth intern. Conf. on Cosmic Rays* **2**, 323 (1977).
24. Fisk, L., Kozlovsky, B. & Ramaty, R. *Astrophys. J.* **190**, L35 (1974).
25. Garcia-Munoz, M., Mason, G. M. & Simpson, J. A. *Astrophys. J.* **217**, 859 (1977).
26. Hagen, F. A., Fisher, A. J. & Ormes, J. F. *Astrophys. J.* **212**, 262 (1977).
27. Parker, E. N. *Stars and Stellar Systems* **7**, ch. 14 (University of Chicago Press, Chicago, 1968).
28. Owens, A. J. & Jokipii, J. R. *Astrophys. J.* **215**, 677 (1977).
29. Cox, D. P. & Smith, B. W. *Astrophys. J. Lett.* **189**, L105 (1974).
30. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148 (1977).
31. Meyer, J. P. *Proc. Fourteenth intern. Conf. on Cosmic Rays* **11**, 3698 (1975).
32. Waddington, C. J. *Proc. Fifteenth intern. Conf. on Cosmic Rays* (in the press).

articles

Possible X-ray counterparts of γ -ray sources

L. Maraschi

Laboratorio di Fisica Cosmica e Tecnologie Relative, CNR, Milano
Istituto di Scienze Fisiche, Università degli Studi, Milano, Italy

T. Markert, K. M. V. Apparao & H. Bradt

Department of Physics and Center for Space Research, Massachusetts Institute of Technology

H. Helmken

Center for Astrophysics, Harvard College Observatory and Smithsonian Astrophysical Observatory

W. Wheaton, W. A. Baity & L. E. Peterson

Physics Department, University of California, San Diego

The error circles for the 10 unidentified COS-B γ -ray sources have been examined for X-ray emission. Three of these sources were found to contain known X-ray sources within their error circles: the positional uncertainty of CG135+1 contains 4U0241+61, CG312-1 contains MX1406-61 and 4U1416-62, and CG327-0 contains 2S1553-542. Spectra and variability, as well as positional coincidence, are used to argue that 4U0241+61, MX1406-61 and 2S1553-542 are possible X-ray counterparts to the γ -ray sources.

RESULTS from the COS-B satellite, which cover 40% of the galactic plane, present evidence for the existence of the 13 localised sources of γ -ray (> 100 MeV) emission shown in Table 1¹. Of these three or possibly four have been previously detected by the SAS-2 satellite²⁻⁴. Three were identified by SAS-2 with objects seen at other wavelengths, namely: CG185-5 with NP0532 (Crab Pulsar), CG263-2 with PSR0833-45 (Vela Pulsar) and CG078-1 tentatively identified with Cyg X-3 because of the 4.8 h pulsations seen in the data⁴. This latter identification is tentative because the pulsations were not confirmed by COS-B (ref. 5) and the position of Cyg X-3 lies just outside the error box for the nearest COS-B source. (However, Cyg X-3 has undergone an intensity transition in X rays since the original SAS-2 observation, which could also have reduced the γ -ray intensity below the COS-B sensitivity threshold.)

The remaining 10 sources are as yet unidentified—there have been no associations established with bright X-ray sources, radio pulsars or supernova remnants. One of these, CG195+4, seen by both SAS-2 and COS-B, seems to exhibit a 59 s period⁵. The latitude distribution of the sources indicates that they are galactic, with a mean distance $\gtrsim 1$ kpc implying luminosities greater than 10^{35} erg s⁻¹ above 100 MeV. These

sources pose extremely interesting astrophysical problems, both as a class, as they can contribute a substantial fraction of the galactic disk emission⁶ and individually, as the emission above 100 MeV appears energetically dominant over the emission in the keV range.

Table 1 COS-B observations of localised γ -ray sources¹

Source	l^{II}	b^{II}	Relative intensity	Remarks
CG 64+0	$64^{\circ}5 \pm 0^{\circ}5$	$0^{\circ}0 \pm 1.0$	$0.30^{+0.10}_{-0.05}$	
CG 75+0	75.0 ± 1.0	0.0 ± 0.5	$0.45^{+0.15}_{-0.10}$	} Could be one extended source CYG X-3?
CG 78+1	78.5 ± 0.5	$+1.5 \pm 0.5$	$1.00^{+0.30}_{-0.20}$	
CG 121+3	121.0 ± 1.0	$+3.5 \pm 1.0$	$0.20^{+0.10}_{-0.05}$	
CG 135+1	135.5 ± 1.0	$+1.5 \pm 1.0$	$0.30^{+0.15}_{-0.10}$	4U0241+61?
CG 176-7	176.0 ± 1.0	-7.0 ± 1.0	0.50 ± 0.10	
CG 185-5	185.3 ± 0.5	-5.6 ± 0.5	1.00*	PSR0531+21
CG 189+1	189.0 ± 1.5	$+1.0 \pm 1.5$	0.40 ± 0.10	Extended?
CG 195+4	195.9 ± 1.0	$+4.5 \pm 0.5$	0.90 ± 0.20	195+5 (SAS-2)
CG 263-2	263.7 ± 0.3	-2.6 ± 0.3	3.30 ± 0.50	PSR0833-45
CG 312-1	312.0 ± 1.0	-1.5 ± 1.0	0.45 ± 0.15	MX1406-61?
CG 327-0	327.5 ± 1.0	-0.5 ± 1.0	0.35 ± 0.15	4U1416-62?
CG 333+0	333.5 ± 1.0	0.0 ± 1.0	0.07 ± 0.2	2S1553-542?

*Absolute intensity of CG 185-5 above 100 MeV is $(3.5 \pm 1.0) \times 10^{-6}$ photons cm⁻² s⁻¹.

Even though the X-ray counterparts of γ -ray sources may be weak, it is essential to perform X-ray searches of the γ -ray error boxes. This is because the γ -ray positions are poorly known and the low statistical precision has limited the study of time variations of the sources. Consequently, an X-ray identification could provide information on temporal behaviour and help identifications at optical and radio wavelength.

Such information is important for clarifying the nature of these objects, for which widely different models have been discussed: (1) Compton scattering on synchrotron photons in radio sources⁷; (2) cosmic rays in supernova remnants⁸; (3) cosmic rays in interstellar clouds⁹; (4) pulsars¹⁰⁻¹⁴; (5) accreting black holes^{15,16}. The observation of fast variability, for instance, would exclude models like (2) and (3) connected with large objects; the association with a binary system would favour (5); a periodic behaviour is likely in (4).

We present here the results of a survey among most of the known X-ray source positions (4 Uhuru¹⁷, M.I.T. OSO-7¹⁸, Ariel 5¹⁹⁻²¹ and M.I.T. SAS-3) for X-ray emitting objects which fall within the error boxes of the 10 unidentified COS-B sources. In three cases, CG135-1, CG312-1, and CG327-0, an X-ray source was found within the γ -ray error box. However, because of the large uncertainty in γ -ray source positions ($1-4^\circ$), the positional coincidence is not necessarily conclusive. In fact, outside the galactic bulge region ($|l| < 60^\circ$) there are seven COS-B sources and ~ 200 X-ray sources within $\pm 15^\circ$ (COS-B scan width) of the galactic plane. This implies that there could be at least one accidental coincidence. Hence, additional information on the spectral or temporal character of the X-ray sources is required to support their identification. A search of the OSO-7 and SAS-3 data has uncovered new information on the spectra and variability of the candidate sources, and is discussed below.

CG135+1

The most probable position for the X-ray source 4U0241+61 = 3U0258+60 lies within the γ -ray error box for CG135-1. 4U0241+61 was observed several times during 1972 by the M.I.T. detectors aboard OSO-7 (ref. 22). Positive detections (greater than 2σ above background) were recorded in the neon, argon and krypton filled proportional counters which covered the ranges 1-6, 3-10 and 15-40 keV respectively. In the 3-10 keV detector, the weighted mean counting rate was 0.34 ± 0.04 counts s^{-1} ($\sim 6.0 \mu Jy$ at 5 keV). This is roughly equivalent to an Uhuru counting rate of 5.1 ± 0.6 counts s^{-1} which, allowing for uncertainties in the conversion from M.I.T. OSO-7 to Uhuru counts, is in reasonable agreement with the value of 3.3 ± 0.5 counts s^{-1} quoted in the Uhuru catalogue¹⁷.

The average spectrum derived from the weighted mean of four observations made between March and December 1972 by OSO-7 is shown in Fig. 1. Also plotted are the 100-MeV flux of CG135+1 from COS-B measurements¹ and two Ariel 5 data points²³, one an upper limit for X rays between 20 and 300 keV and one positive detection of low-energy γ -rays between 0.3 and 1.2 MeV made by the scintillation detectors on Ariel 5. As indicated in Fig. 1, the X-ray spectrum of 4U0241+61 is quite hard (the power law index of the energy spectrum is consistent with zero) and may remain flat out to, or beyond 1 MeV.

As discussed above, the positional coincidence of CG135+1 and 4U0241+61 is not sufficient to indicate a physical connection between the sources. On the other hand, when the positional coincidence and the hard X-ray and low-energy γ -ray spectrum observed by OSO-7 and Ariel 5 are taken together, then the physical association becomes more convincing. It is possible, therefore, that 4U0241+61 is the X-ray counterpart of CG235+1.

CG312-1

This γ -ray source falls in a complex X-ray emission region. A transient source²⁴ was observed there above 7 keV by the

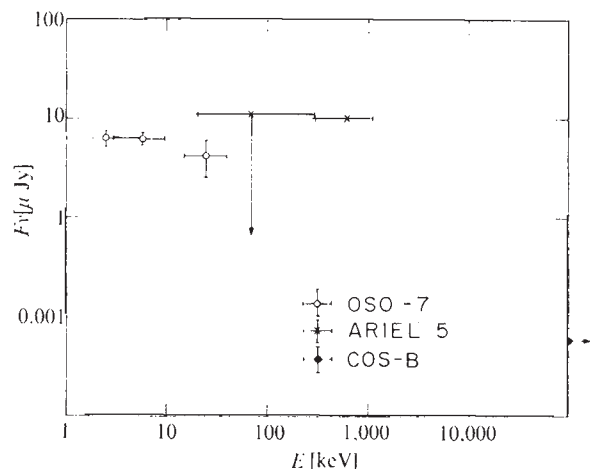


Fig. 1 Spectrum derived from OSO-7 (ref. 16), Ariel 5 (ref. 23), and COS-B (ref. 1) observations of 4U0241+61/CG135+1. No error bars were quoted for the Ariel 5 positive detection (0.2-1.2 MeV).

UCSD hard X-ray telescope on OSO-7 during the spring of 1972, and the M.I.T. OSO-7 detectors showed no less than 3-10 keV sources within the 3° radius UCSD error circle at about the same time, MX1353-64 (= A1354-64), MX1406-61, and MX1418-61 (= A1415-60?, = 4U1425-61?)²⁵. The 4 Uhuru catalogue¹⁵ lists 4U1416-62, 4U1425-61 (= A1415-60? = MX1418-61?) and 4U1344-60 (= A1343-60) in the same area. Subsequent refinement of the UCSD data has reduced the size of the error circle but there are still at least four sources within or near it (Fig. 2). The UCSD position, MX1406-61 and 4U1416-62 are within the error box for CG312-1 and MX1418-61 lies only $\sim 0.5^\circ$ outside. Figure 2 displays these and several other OSO-7, Uhuru and Ariel 5 sources.

The UCSD instrument²⁶, a 64-cm² area NaI scintillation telescope, actively collimated to a circular 6.5° (FWHM) field of view, and sensitive in the 7-500 keV range, showed less than 2×10^{-2} photons $cm^{-2} s^{-1}$ ($\lesssim 10 \mu Jy$) from 7-55 keV from the region of CG312-1 during a scan in the plane of the Galaxy

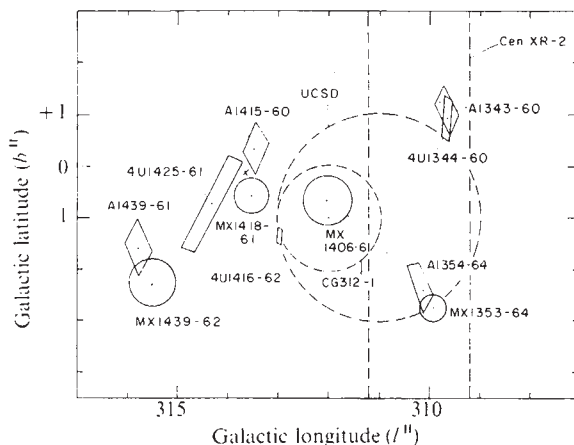


Fig. 2 Summary of X-ray and γ -ray observations of the region near $(l, b) = (312, 0)$ by Uhuru (ref. 17), OSO-7 (ref. 18), Ariel-5 (ref. 21) and COS-B (ref. 1). The UCSD circle was previously unpublished. The X between 4U1425-61, MX1418-61 and A1415-60 is the position assumed for MX1418-61 and for the purposes of deriving the M.I.T. spectra shown in Fig. 3. Slightly different assumptions for the position of MX1418-61 are reflected in somewhat different derived spectra but give a less satisfactory fit to the data.

in late December 1971. Between 25 February and 2 March 1972, the 7–28 keV flux was $(1.9 \pm 0.63) \times 10^{-2}$ photons $\text{cm}^{-2} \text{s}^{-1}$ ($\sim 10 \mu\text{Jy}$) and reached a maximum of $(13 \pm 0.6) \times 10^{-2}$ photons $\text{cm}^{-2} \text{s}^{-1}$ ($\sim 65 \mu\text{Jy}$) between 15 and 23 March 1972. The spectrum at this and later times is shown in Fig. 3. The initial post-maximum decline left the spectrum unchanged. However, between 8 and 23 June 1972 the source brightened to its previous high intensity from 50 to 80 keV, while no flux was detected below 20 keV. This episode climaxed on 21 June 1972, with a flare of 3σ significance, during which the hard X-ray flux increased by a factor of 4 in less than an hour, and then decayed to its original level over the next six hours.

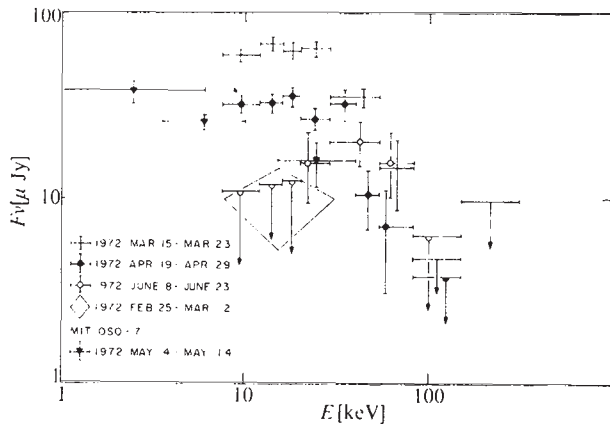


Fig. 3 M.I.T. and UCSD OSO-7 observations of the spectrum of the 1972 Centaurus transient in the region of CG312–1. The UCSD analysis assumes the position of the source is $l^{\text{II}} = 311^\circ$, $b^{\text{II}} = -1^\circ$, as best fits the hard X-ray data. Reasonable alterations of this number could change the overall normalisation by up to 30%, without much effect on the spectral shape. The M.I.T. data are for the source MX1406–61. If this graph were extended, the COS-B data point for CG312–1 would lie at $> 10^2$ keV and $F_v = (9 \pm 3) \times 10^{-3} \mu\text{Jy}$.

The UCSD data are consistent with the position ($l^{\text{II}} = 310.2 \pm 1.0^\circ$) of the April 1967 transient Cen XR–2²⁷, as are the position of MX1353–64 and 4U1344–60. However, spectral information from the 3° (FWHM) field of view M.I.T. detector weighs against such an identification, supporting MX1406–61 as the probable hard source. The M.I.T. instrument scanned the region from 4 to 14 May 1972. A unique interpretation of the data is obscured by the presence of three sources, at least one of which is known to be variable (Fig. 2). However, the spectrum (energy power law index $\alpha \sim 0.5$) and intensity of MX1406–61 from the M.I.T. data and the hard X-ray results reported above are consistent with each other when extrapolated to the same energy (Fig. 3).

The spectrum described above is similar to those of hard X-ray transients like A0535–26 (ref. 28) or to the spectra of some of the accretion-powered binaries²⁹, although the occurrence of a hard X-ray flare is less common. On the basis of spectrum, brightness and variability, therefore, we believe MX1406–61 to be a likely counterpart of CG312–1. We emphasise, however, that on the basis of positional coincidence alone, 4U1416–62 is an equally good candidate. A deep survey of the region might detect either or both of these sources.

CG327–0

The X-ray source 2S1553–542^{30,31} lies within the error box of CG327–0. The X-ray object, discovered by the SAS-3 satellite, had a 2–11 keV intensity of $27 \mu\text{Jy}$ or about

25 Uhuru counts s^{-1} (2–6 keV). The failure of Uhuru to report this source, therefore, suggests that it is variable. There is also some evidence that it possesses a hard spectrum. The SAS–3, 2–11 keV data are consistent with a nearly flat ($\alpha \sim 0.5$) spectrum. There is no optical identification for 2S1553–542. A B0 supergiant and the variable star TT Normae are within $2'$ of the X-ray position but well outside the $35''$ radius error circle.

Conclusions

Of the 13 COS-B γ -ray sources, six have been linked to X-ray sources (including the three discussed here). The SAS-3 rotating modulation collimator data were examined to determine upper limits for the COS-B sources not discussed above. For each of the seven error regions we place upper limits of $10 \mu\text{Jy}$ (~ 6 Uhuru counts s^{-1}) on X-ray emission by a point source in the 2–11 keV range.

The results cited in this paper indicate that the X-ray source 4U02416+1 is a possible candidate as the X-ray counterpart of CG135+1, both because of positional coincidence and spectral hardness. A significantly improved source position (fairly uncrowded region of the X-ray sky) and optical counterpart identification are needed. The existing measurements also demonstrate a hard X-ray and low energy γ -ray flux which should be detectable by balloon and satellite experiments. Any γ -ray, X-ray, optical or radio data should be examined for pulsations and a light curve history which would yield the overall variability of the source.

The identifications of CG312–1 and GC327–0 with MX1406–61 and 2S1553–542 respectively cannot be made with great confidence. For MX1406–61 the spectrum, particularly the UCSD data, suggests a hard X-ray component, but source confusion in the region near $l^{\text{II}} = 312^\circ$ makes identification uncertain. For 2S1553–542, there is less source confusion but the spectral data are incomplete. Improved positional information (leading to optical or radio identifications) and spectra, particularly at high energies, could confirm or disprove these tentative associations.

This work was supported in part by NASA contracts NAS 5-11080, NAS 5-11082, NAS 5-11450. The UCSD group acknowledges the analysis support of M. Ulmer. L. Maraschi thanks M.I.T. for the hospitality during her visit.

Received 12 December 1977; accepted 19 January 1978.

- Hermesen, W. *et al.* *Nature* **269**, 494 (1977).
- Thomson, D. J., Fichtel, C. E., Kniffen, D. A. & Ogelman, H. B. *Astrophys. J.* **200**, L79 (1975).
- Thomson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252 (1977).
- Lamb, R. C., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Thomson, D. J. *Astrophys. J.* **212**, L63 (1977).
- Masnou, J. L. *et al.* *Proc. 12th ESLAB Symp.*, Frascati, Italy, ESA SP124 (1977).
- Biggami, G. F., Caraveo, P. & Maraschi, L. *Astr. Astrophys.* (in the press).
- Gould, R. J. *Phys. Rev. Lett.* **15**, 577 (1965).
- Pinkau, K. *Phys. Rev. Lett.* **25**, 603 (1970).
- Black, J. H. & Fazio, G. G. *Astrophys. J.* **185**, L7 (1973).
- Highdon, J. C. & Lingenfelter, R. E. *Astrophys. J.* **208**, L107 (1976).
- Cheng, A. F. & Ruderman, M. A. *Astrophys. J.* **216**, 865 (1977).
- Hardee, P. H. *Astrophys. J.* **216**, 873 (1977).
- Ozerov, L. M. & Usov, V. V. (personal communication).
- Massaro, E. & Salvati, M. *Proc. 12th ESLAB Symp.*, Frascati, Italy, ESA SP124 (1977).
- Illarionov, A. F. & Sunyaev, R. A. *Astr. Astrophys.* **39**, 185 (1975).
- Maraschi, L. & Treves, A. *Astrophys. J.* **218**, L113 (1977).
- Forman, W. *et al.* CFA preprint no. 763 (1977).
- Markert, T. H. *et al.* preprint CSR-P-77-15 (1977).
- Seward, F. D., Page, C. G., Turner, M. J. L. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **175**, 39P (1976).
- Seward, F. D., Page, C. G., Turner, M. J. L. & Pounds, K. A. *Mon. Not. R. astr. Soc.* **177**, 13P (1976).
- Cooke, B. A. *Mon. Not. R. astr. Soc.* **182**, 489 (1978).
- Clark, G. W., Bradt, H. V., Lewin, W. H. G., Markert, T. H., Schnopper, H. W. & Sprott, G. F. *Astrophys. J.* **179**, 263 (1973).
- Coe, M. J., Engel, A. R. & Quenby, J. J. *IAU Circ.* No. 3003 (1976).
- Wheaton, W. H., Baity, W. A. & Peterson, Q. E. *IAU Circ.* No. 2761 (1975).
- Markert, T. H. *et al.* *Astrophys. J.* (in the press).
- Peterson, L. E. *A. Rev. astr. Astrophys.* **15**, 213 (1975).
- Francey, R. J. *Nature phys. Sci.* **229**, 229 (1971).
- Maraschi, L., Huckle, H. E., Ives, J. C. & Sanford, P. W. *Nature* **263**, 34 (1976).
- Ulmer, M. P. *Astrophys. J.* **196**, 827 (1975).
- Apparao, K. M. V., Bradt, H. V., Dossay, R., Dower, R., Jernigan, J. G. & Li, F. K. *Nature* **271**, 225–228 (1978).
- Walter, F. *IAU Circ.* No. 2959 (1976).

Geos I: Identification of natural magnetospheric emissions

P. Christiansen, P. Gough & G. Martelli

Plasma and Space Physics Group, School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton, Sussex, UK

J.-J. Bloch, N. Cornilleau, J. Etcheto & R. Gendrin

Centre de Recherches en Physique de l'Environnement Terrestre et Planétaire, CNET, Issy-les-Moulineaux, France

D. Jones

Space Science Department, European Space Agency, ESTEC, Noordwijk, Netherlands

C. Béghin & P. Décréau

Centre de Recherches en Physique de l'Environnement Terrestre et Planétaire, CNRS, Orleans-la-Source, France

Natural noise bands, a persistent feature of the Earth's magnetosphere, have been identified for the first time, using novel techniques on the recently launched Geos I satellite. It is argued that these waves, in the electron cyclotron harmonic mode, are radiated incoherently by suprathermal electrons.

GEOS I, the European Space Agency's magnetospheric research satellite, was launched on 20 April 1977 from the Eastern Test Range, US. Because of a failure of the Delta launcher, the satellite did not achieve the intended geostationary orbit, and was subsequently manoeuvred into a 12 h (perigee 2,000 km, apogee 40,000 km ($7.1 R_E$)). The research proposed originally has thus been severely limited as the satellite spends only about five hours per orbit around the geostationary position, and the slow drift of orbit elements means that a full year is needed to sample all local times. Nevertheless, new and interesting results have been and are being recorded. Here we present data which allow us to answer, with reasonable precision, several questions about electrostatic wave emissions observed by earlier spacecraft. It should be stressed that these results are only one part of a larger set of observations of various kinds of wave phenomena which are

being studied. We shall concentrate on the rather weak emissions which show a pronounced banded structure close to the local plasma frequency.

Two facilities provided by the Geos wave experimental package have been essential for this particular study¹: (1) the 'passive' natural wave detector, and (2) the two 'active' wave experiments, which provide important diagnostic measurements of the basic plasma parameters. Both require a brief description.

Passive wave experiment

The natural wave detection experiment is made up of several units, which may be connected to a number of wave sensors. The data presented here come exclusively from the large electric antenna, which consists of a pair of spherical probes containing sensitive preamplifiers mounted at the end of 20-m booms and deployed perpendicular to the satellite's spin axis (the y -axis in spacecraft coordinates). We are only concerned with the signals from this sensor which are sampled by a stepped frequency analyser (SFA) unit. The SFA operates by telecommand and can be programmed in 300-Hz steps from any minimum frequency to almost any maximum frequency up to 77 kHz. This flexibility of command is a powerful operating feature of Geos satellite.

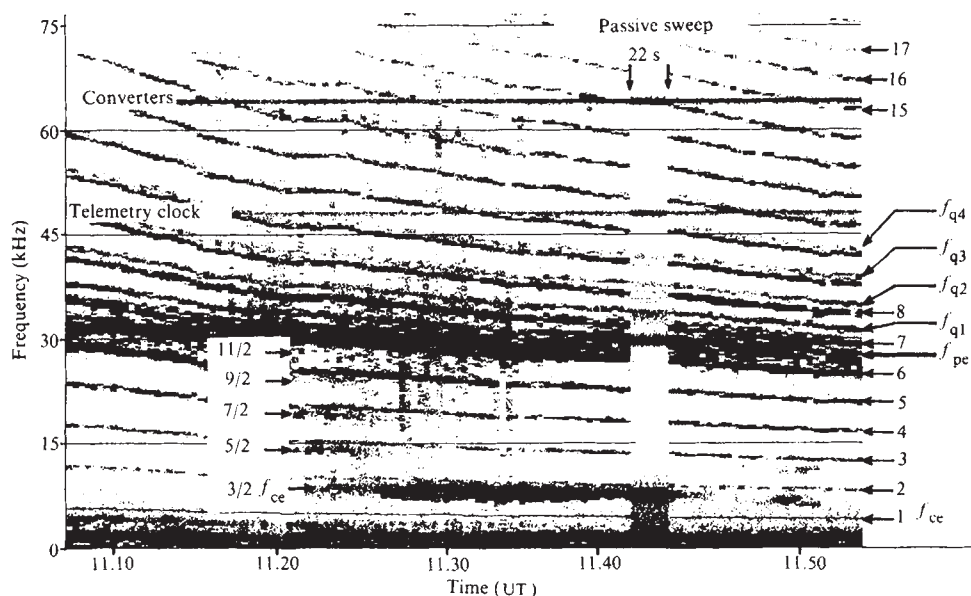


Fig. 1 Spectrogram of active sounder results showing gyroharmonics, f_p and f_q resonances. Matching natural emissions at $\sim f_p - f_{uh}$ and lower f_q frequencies are seen in the time expanded passive sweep at 11.41. Other natural signals in the whistler mode below f_{ce} , and between lower gyroharmonics can also be seen. The grey scaling, black to white = 20 db is set to show general spectral detail, the grey level being set for each 22-s frequency sweep to the spectral mean.

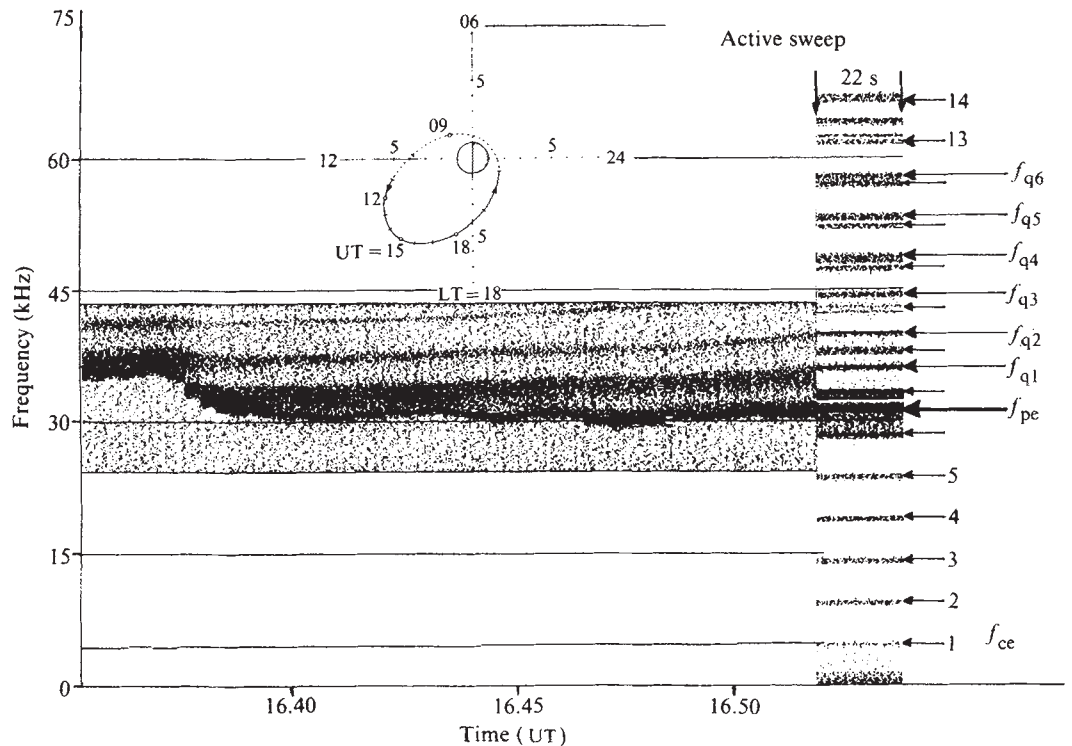


Fig. 2 Spectrogram of natural emissions recorded by the SFA sweeping around f_{pe} . The natural bands match the $f_{pe} - f_{uh}$ and f_q resonances shown on the active sweep at the end of the display. (Inset: satellite trajectory from which local times of observations in Figs 1 and 2 can be inferred).

Active wave experiments

The two active diagnostic experiments are a plasma resonance sounder² and a mutual impedance probe. The resonance sounder works as follows. A 3-ms burst of sinusoidal signal, with amplitude ~ 100 V is applied to the 2×20 m dipole formed by the boom stems of the long electric antenna. At the end of the emission, the SFA receiver with an input bandwidth matched to the transmitted pulse, is switched to the large electric antenna and the temporal decay of the pulse is monitored for ~ 80 ms. The pulse and receiver frequency are then varied in 300-Hz steps over the full 77-kHz range of the SFA. A total sweep lasts ~ 22 s.

Resonances are observed during the monitoring period when the pulsed emission contains frequencies that couple to weakly damped, low group velocity wave modes in the surrounding plasma. In a hot magnetised plasma the resonance frequencies include the electron gyro frequency $f_{ce} = Be/2\pi m$ and its harmonics (nf_{ce}), the electron plasma frequency $f_p = ne^2/2\pi\epsilon_0 m$, and the f_q resonances. All these resonances can be understood in terms of the properties of normal wave modes in a plasma, and the f_q resonances in particular are defined by the dispersion properties of electron cyclotron harmonic (ECH) waves, often called Bernstein waves³.

For the simplest case, where the wave number k is exactly perpendicular to the ambient magnetic field B_0 , these waves propagate without damping and, for a Maxwellian plasma, they have the following dispersion relation

$$D(\omega, k) = 1 + \omega_p^2/k^2 v_t^2 \left[1 - \sum_{s=-\infty}^{\infty} I_s(\lambda) e^{-\lambda} (\omega/(\omega - s\omega_c)) \right] = 0 \quad (1)$$

where $\lambda = kv_t/\omega_c$, v_t is the electron thermal speed and I_s is the modified Bessel function of order s .

It can be seen that above the upper hybrid frequency $f_{uh} = (f_p^2 + f_{ce}^2)^{1/2}$ there are frequencies lying between successive cyclotron harmonics at which the group velocity $v_g = (\partial\omega/\partial k) = 0$. These are the f_q resonances, which are shown, for a particular choice of plasma parameters, in Fig. 4b. Given that these resonances can be quickly identified, the density and magnetic field can be measured accurately or deduced through their functional dependence in equation (1).

The mutual impedance probe measures the (frequency dependent) mutual impedance $Z(\omega)$ between a transmitting dipole,

formed by two of the short axial electric probes, and a receiving dipole, the large electric antenna. In this configuration $Z(\omega)$ is basically determined by the dispersion of electron plasma waves parallel to B_0 , that is, by f_p and the Debye length $\lambda_D = v_t/\omega_p$. This allows the plasma density and temperature to be measured.

However, as the dipoles are of unequal lengths and as the line between transmitter and receiver is not always parallel to B_0 , additional contributions to $Z(\omega)$ can be expected from waves propagating obliquely to the magnetic field.

Experimental results

Before discussing specific examples we should broadly categorise the frequencies at which signals may occur in this experiment. There are two classes depending on whether active or passive measurements are made. These are not independent because the excited waves obey the same dispersion relations, but as the excitation mechanisms are quite dissimilar (active perturbations and plasma electrons, respectively) different spectral detail is to be expected. (1) For the active measurements, moving up from low frequencies, resonances occur at harmonics of f_{ce} , then at the local plasma and upper hybrid frequencies, and finally at the f_q frequencies, as discussed above. (2) The occurrence of natural waves depends sensitively on the precise electron-velocity distribution, but electrostatic waves signals may occur (a) between low harmonics of f_{ce} (ECH waves), (b) close to f_{pe} (plasma waves) and (c) close to f_{uh} and the f_q frequencies (ECH waves).

Details of (a), (b) together with electromagnetic waves propagating above f_{pe} and electromagnetic and ion waves below f_{ce} will be published elsewhere.

The most persistent signals observed in the first six months of Geos operation are weak banded emissions. They occur close to what was identified, by experimenters concerned with both diagnostic and natural waves, as the local plasma frequency. These emissions are only detected by the long electric antenna (sensitivity $\sim 10^{-8}$ V m⁻¹ Hz^{-1/2}) and typically have field amplitudes of a few μ V m⁻¹ in a 300-Hz band. They are seen for many hours during a satellite pass, essentially whenever f_{pe} is below the maximum observable frequency of 77 kHz. Signals of this kind are noted in data from the 1 MP-6 satellite⁴, but because precise plasma diagnostics were lacking, were not identified.

The results presented below are examples of a detailed investigation of these emissions, in which the flexible Geos telecom-

mand system is used to mix natural wave detection and diagnostic measurements. Data recorded on 14 July 1977 are shown in Figs 1

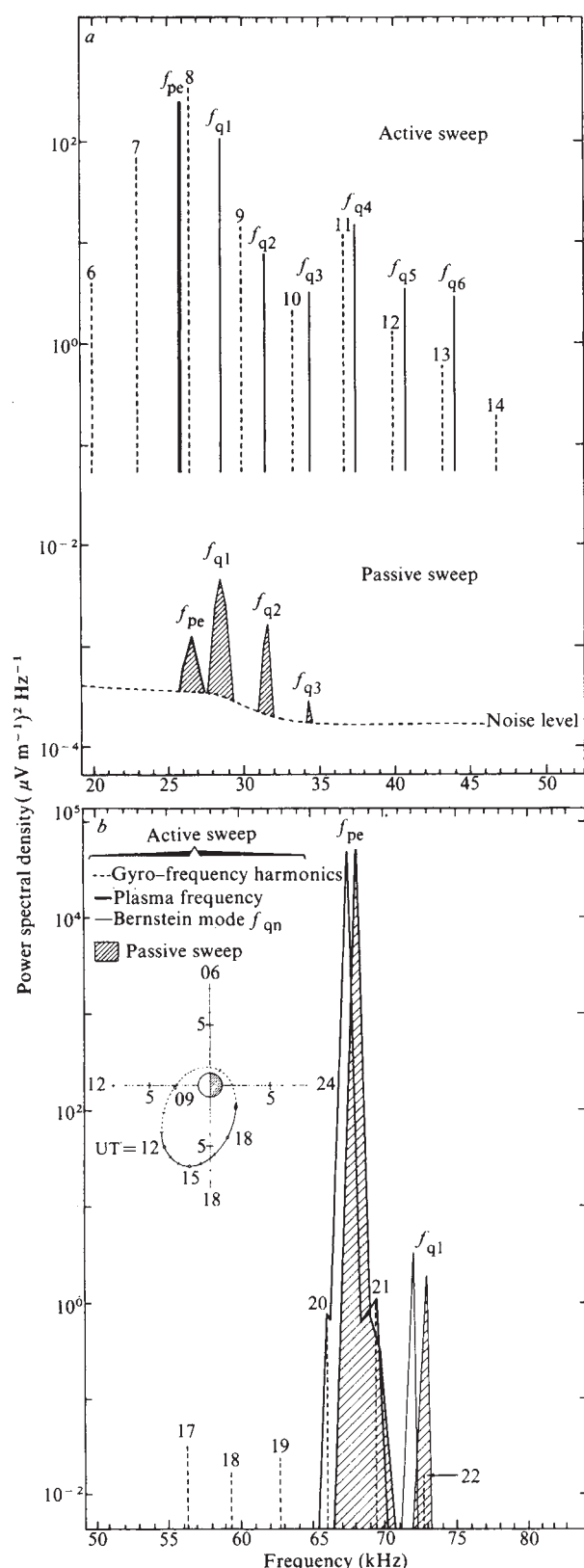


Fig. 3 *a*, Matching of f_p - f_{uh} , f_q frequencies and power levels shown in detailed comparison of a natural bottom, and active power spectrum recorded 22 s later where $f_{ce} = 3.33$ kHz. The dashed lines show level, frequency and number of gyroresonances (15 June 1977, 16.03 UT). *b*, An example of strong natural signals close to f_{pe} , shown on successive passive and active frequency sweeps (17 June 1977, 12.03 UT).

and 2. In Fig. 1 we show the time development of frequency spectra produced by the resonance sounder, interspersed with passive spectra, the vertical frequency axis covers the full 0.77 kHz range of the SFA. The active spectra beginning at 11.07 UT show a multiplicity of resonances. Starting from low frequency, a series of features at harmonics of f_{ce} can be seen, in agreement with the spacecraft magnetometer measurement which is shown as a continuous line beginning at ~ 6 kHz. This ascending series is broken by a broader, more complex feature which is identified as the plasma frequency. Both the f_p resonance and the nf_c series show a general decrease in frequency as the satellite moves to higher altitudes. Above f_p more cyclotron harmonics can be seen, now accompanied by additional resonances whose separation from nf_c decreases with increasing frequency. That these are correctly identified as f_q resonances will be seen later. Spectra of natural signals, recorded at 11.42 and shown on an expanded time scale, reveal a triplet of emissions. The strongest occurs close to f_p (and also to f_{uh} which exceeds f_p by only a few percent for $f_p/f_{ce} \sim 6$), while the other two are rather close to the first couple of f_q resonances. Note also the appearance of a natural emission beginning at ~ 11.22 between f_c and $2f_c$ which is strong enough to be recorded during the sounder monitoring period. This is an example of the so called $(n + \frac{1}{2})f_{ce}$ emission⁵, also regularly observed on Geos 1.

Data from the same day, but with a different emphasis, are shown in Fig. 2. The plasma frequency has been estimated in real time from an earlier active measurement, and the SFA is commanded to sweep the frequency range around it. The display terminates with a time expanded spectrum from the active sounder. Once again a triplet of natural frequencies can be seen, which show interesting temporal detail, corresponding rather closely to the values of f_p/f_{uh} and the two nearby f_q resonances shown in the active cycle.

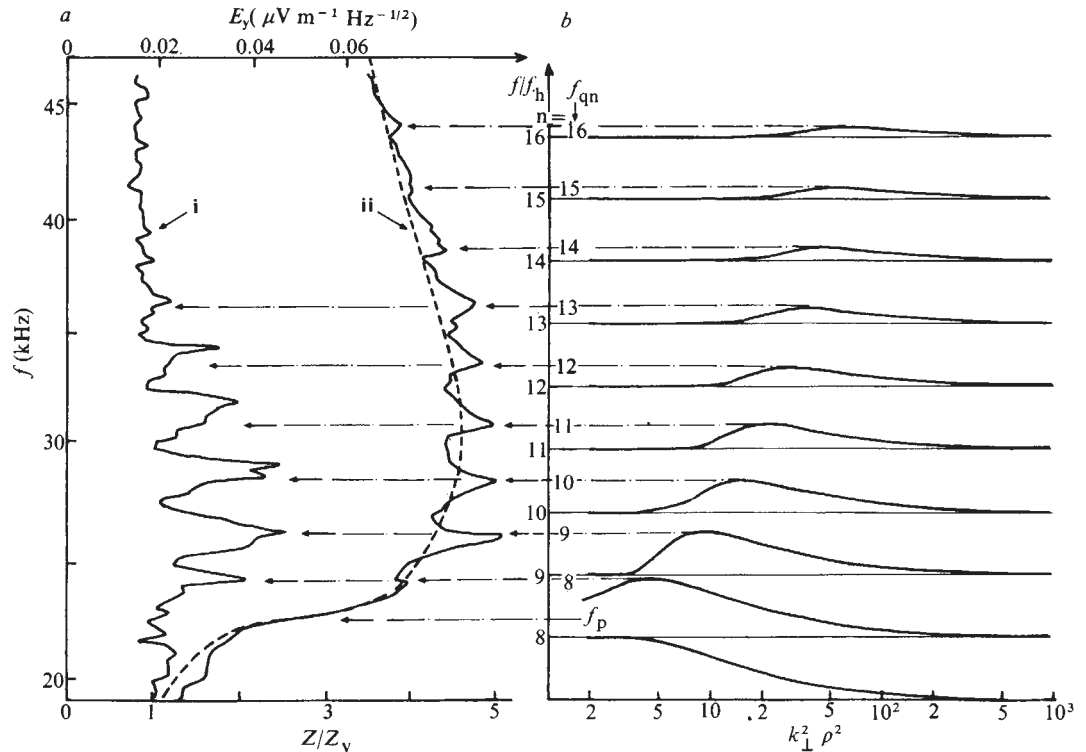
The remaining figures provide further corroborative evidence from different days, by comparing passive spectra with the results of active measurements. Figure 3*a* shows a single power spectrum of natural waves, and on the same scale, shown as vertical lines, are the frequencies and power levels of active sounder resonances recorded 22 s later. A fourth band can be clearly identified in the natural spectrum and the f_p emission power is lower than that of the adjacent f_q , whereas the opposite is true in Figs 1 and 2. This sort of variability in detail is typical for these emissions, but the identification of frequencies remains clear. Figure 3*b*, however, demonstrates another phenomenon. Now there is very little difference in the power level of the f_p lines recorded on active and passive sweeps and the levels are extremely high, ~ 65 db above the quiet spectrum for Fig. 3*a*, and ~ 30 db above even the active resonance in the same figure. This is due to the onset of a natural emission, seen occasionally on Geos, of sufficient strength to obscure any f_p resonances caused by the sounder, and depress nearly all cyclotron resonances. The single f_q signal that does appear is also believed to be of natural origin, and the displacement of two spectra simply a result of rapid variations in the plasma density on a time scale of 22 s.

Figure 4, made early in the satellite life when the long booms were only half deployed, contains a comparison of the natural emissions with the result of a mutual impedance measurement. This time, the passive spectrum contains even more detail and five f_q values can be seen. In Fig. 4*a*, the mutual impedance $Z(\omega)$ displays clearly matching spectral detail superimposed on the general form of the field free impedance function, bearing out our contention that oblique wave modes contribute to the impedance function. Figure 4*b* is the result of a theoretical fit of equation (1) to the data. The values of f_p and electron temperature are inferred from the basic impedance curve, and the remaining free parameter, f_{ce} , is varied to best fit the observed f_q frequencies. The value of f_{ce} thus found is within 1% of that recorded by the spacecraft magnetometer.

Discussion

The close correspondence between the frequencies of the weak natural emissions and the plasma/upper hybrid, and nearby f_q

Fig. 4 Identification of f_q resonances by the mutual impedance experiment, on 10 May 1977, at 18.00 UT *a*, Curve i shows the electric field E_y measured against frequency f during a passive sweep. Curve ii shows the normalised mutual impedance, Z/Z_v , against frequency obtained 22 s later. The experimental response (solid line) is given together with the best fit theoretical response in the isotropic approximation (dashed line) ($n_e = 6.25$ electrons cm^{-3} , and $T_e = 13,500$ K) *b*, Dispersion diagram for a perpendicular ECH wave propagation. The frequency normalised to f_c (here: 2.722 kHz) is given against wave number $k_\perp$ normalised to the Larmor radius ρ . Arrows join the frequencies of f_q defined in the right-hand diagram, to the peaks of the experimental left-hand curve.



resonance frequencies derived from diagnostic measurements, is powerful evidence that the waves are propagating in the ECH wave mode, as these have the observed banded frequency structure and are known to be marginally stable (undamped) for purely perpendicular propagation. For the more general case of oblique propagation ($k_\parallel \neq 0$), the dispersion relationship $D(\omega, k_\perp, k_\parallel)$ is more complex than that shown in Fig. 4b (ref. 6). Assuming an isotropic Maxwellian distribution $f(v)$, (which is certainly not true for magnetosphere) a given choice of $\omega, k_\perp$ now gives two values of $k_\parallel$. The wave numbers are now complex, indicating wave damping, which increases as the propagation angle departs further from perpendicularity. The two branches are further modified if the assumption of isotropy is relaxed.

In a plasma, radiation occurs in these modes by the transfer of energy from electrons obeying the condition

$$\omega - k_\parallel v_\parallel = \pm s\omega_c$$

which contains both Cerenkov ($s = 0$) and cyclotron ($s = 1, 2, \dots$) resonances. The energy in these oscillations increases significantly above the thermal level in two circumstances. First, the ECH wave dispersion relation contains growing solutions if $f(v)$ is strongly anisotropic or contains regions of velocity space where $\partial f / \partial v > 0$ (ref. 6). These coherent instabilities generally reach large amplitudes, and are stabilised by quasi-linear or nonlinear processes.

The second mechanism invokes incoherent processes, in which a rather lower level of fluctuations is determined by the sum of contributions from single 'test' electrons radiating by Cerenkov or cyclotron emission⁷. For these waves not to be quickly reabsorbed, the radiating particles must lie well away from the body of the velocity distribution. This can occur in plasma containing a near Maxwellian component and a small population of non-thermal high energy electrons.

We favour the incoherent mechanism as an explanation for the weak emissions for two reasons. First, it is most unlikely that the electron velocity distribution could maintain the level of anisotropy required for coherent emission for the long periods (~ 6 h) observed, under varying conditions of density and magnetic field during a satellite pass. Second, the example in Fig. 3b shows that extremely strong sporadic signals near f_{pe} can exist, and the relative ratio of 65 db between these and the f_q emissions gives the sort of contrast that might be expected between coherent and

incoherent processes.

The theoretical power loss of a single, 100 eV electron has been examined by Trulsen⁸ in a complex calculation. The plasma parameters he uses ($f_p = 2.3 \times 10^5$, $f_c = 10^5$ Hz) are more appropriate to the plasmasphere than those encountered along the Geos orbit, but we regard it as significant, in the light of Figs 1 and 2, that the theoretical power spectra show similar features to our data, that is, an enhanced band which stretches from f_p to above f_{uh} , and weaker bands above this between cyclotron harmonics and their companion f_q frequencies. These bands correspond to frequencies at which low group velocity weakly damped waves exist, and are not expected to vary much for weakly anisotropic background distributions, although their relative amplitudes will depend on the separation of f_{uh} and the nearest gyro harmonic.

Starting from the observation that the f_q emissions observed on Geos are radiated locally in the ECH mode, we can estimate roughly the expected incoherent power level using Trulsen's calculation as a basis. The satellite antenna samples radiation from the particles in a volume of scale length L , which is determined by the damping length of the waves. If the radiated power per particle per unit frequency is $P(\omega)$ then the total power flux per Hz at the satellite is

$$\phi = n' P(\omega) L / 2 = \epsilon_0 E^2(\omega) \langle v_g \rangle / 2$$

where n' is the number density of radiating electrons, $\langle v_g \rangle$ is the average group velocity, and $E(\omega)$ the electric field strength in $\text{Vm}^{-1} \text{Hz}^{-1/2}$. Using the following: $E^2(\omega) \sim 10^{-14} \text{V}^2 \text{m}^{-2} \text{Hz}^{-1}$, $v_i \sim 10^6 \text{ms}^{-1}$ (typical measured values); $\langle v_g \rangle = 0.2v_i$, $P(\omega) = 10^{-27} \text{W Hz}^{-1}$ (estimated from the theory) $n' = 10^3 \text{m}^{-3}$ (estimated from preliminary particle data for 500 eV electron fluxes), the resulting value for L is 20 km, corresponding to a reasonable 40 wavelengths when compared with the theory. This apparently satisfactory result is subject to considerable uncertainty and in particular we believe that $P(\omega)$ may be overestimated but partly compensated by the low estimate of n' . A detailed calculation using observed particle distributions and more realistic plasma parameters is obviously necessary.

Conclusion

With the aid of on-board plasma diagnostic experiments, natural emission bands have been identified with the local plasma-upper

hybrid and nearby f_q frequencies, with consequent and obvious diagnostic implications. The coexistence of the banded, and $(n + \frac{1}{2})f_{ce}$ signals shows that the classification of magnetospheric instabilities is more complex than the current conjecture⁴ that both emissions are basically the same phenomenon, occurring in different regions of the magnetosphere.

We thank M. Petit, B. de la Porte, R. Grard, A. Pedersen, M. Debie, F. Lefeuvre, M. Hamelin, A. Bahnsen, E. Ungstrupp, G. Wrenn, H. Borg and F. Mariani for their valuable scientific collaboration. The contributions of the European Space Operations Centre, the Centre National d'Etudes Spatiales, Toulouse, E. Gershuny and J. Hunter of Sussex University to the data

handling are acknowledged, as is the work of K. Knott, project Scientist, and the Geos project team. This research is supported by the SRC, CNET, CNRS, France, and ESA.

Received 29 March; accepted 30 March 1978.

1. Knott, K. *ESRO Rep. 17th Cospar Ming.* Sao Paulo, Brazil, 162 (1974).
2. Etcheto, J. & Petit, M. *C. r. heb. Acad. Sci. Paris*, **285B**, 329 (1977).
3. Bernstein, I. B. *Phys. Rev.* **109**, 10 (1958).
4. Shaw, R. & Gurnett, D. A. *J. geophys. Res.* **80**, 4259 (1975).
5. Kennel, C. F., Scarf, F. L., Fredericks, R. W., McGhee, J. H. & Coronite, F. V. *J. geophys. Res.* **75**, 6136 (1970).
6. Tataronis, J. A. & Crawford, F. W. *J. Plasma Phys.* **4**, 231, 249 (1970).
7. Biskamp, G. *Radiation Process in Plasmas* (Wiley, New York, 1966).
8. Trulsen, J. J. *Plasma Phys.* **6**, 367 (1971).

Image reconstruction from incomplete and noisy data

S. F. Gull & G. J. Daniell*

Mullard Radio Astronomy Observatory, Cavendish Laboratory, Madingley Road, Cambridge, UK

Results are presented of a powerful technique for image reconstruction by a maximum entropy method, which is sufficiently fast to be useful for large and complicated images. Although our examples are taken from the fields of radio and X-ray astronomy, the technique is immediately applicable in spectroscopy, electron microscopy, X-ray crystallography, geophysics and virtually any type of optical image processing. Applied to radioastronomical data, the algorithm reveals details not seen by conventional analysis, but which are known to exist.

In many branches of physics, such as spectroscopy, X-ray crystallography and radioastronomy, the primary data are the Fourier transform of the physical quantities which are of interest. The inverse Fourier transform is normally then calculated with a digital computer. It is perhaps unfortunate that this calculation is so simple and fast, because it encourages the belief that this is the best way to use the data. However, because the data can never in practice be obtained for the whole range of the transform variable, and can never be free from experimental errors, other possibilities need to be explored.

We have used an alternative method of analysis, based on the idea of maximum entropy, which is illustrated using data produced by the radio telescopes at the Mullard Radio Astronomy Observatory. In this application the data are samples of the two-dimensional Fourier transform of the radio brightness from the sky. We then show how the method may be modified to deal with data not involving a Fourier transform, and apply it to an X-ray map of the supernova remnant Cassiopeia A.

Safe inversion techniques

The correct formulation of problems involving incomplete and noisy data lies in the field of inverse theory. Previously, inverse problems have frequently been tackled by fitting empirical models, although this often leads to false confidence in the conclusions obtained. A notable advance is the method of Backus and Gilbert, which is most clearly explained in the general review by Parker¹. This method shows explicitly the inevitable compromise that must be made between spatial resolution and the accuracy of amplitude measurements, and protects the user against making deductions that are not warranted by the data. This method has been applied to the evaluation of Fourier transforms by Oldenburg².

To avoid abstraction, we shall refer to our radioastronomical example. Starting with incomplete and noisy data, one can obtain by the Backus–Gilbert method a series of maps of the distribution of radio brightness across the sky, all of which are consistent with the data, but have different resolutions and noise levels. From the data alone, there is no reason to prefer any one of these maps, and the observer may select the most appropriate one to answer any specific question. Hence, the method cannot produce a unique 'best' map of the sky. There is no single map that is equally suitable for discussing both accurate flux measurements and source positions.

Nevertheless, it is useful to have a single general-purpose map of the sky, and the maximum-entropy map described here fulfils this role. Although it cannot be thought of as the true sky convolved with any beamshape, we believe that the maximum-entropy map may be used with confidence to describe the general appearance of the sky and that, like the Backus–Gilbert maps, it cannot lead to conclusions for which there is evidence in the data.

The advantages of the maximum-entropy map are: that the radio intensity is shown as everywhere positive, in spite of noise in the data; that sidelobes are automatically reduced to a very low level; the maximum resolution consistent with the physical size of the telescope is obtainable; that the resulting map is smooth; that regions of low surface brightness are clearly shown and are not confused by noise; the method is applicable to the cases where the phases of the Fourier components are either inaccurate or entirely unknown.

The maximum-entropy solution

The origins of our method are given elsewhere^{3–6}. We shall not give quasi-philosophical arguments concerning maximum entropy in this context, but the following discussion shows in physical terms what our mathematics describes. Although we discuss the maximum-entropy method in the radioastronomical context it will be apparent how the method can be applied in other circumstances.

As in statistical mechanics, we introduce quantisation to simplify the counting of states, so we divide the sky into small areas, and quantise the radio intensity. We now imagine the traditional 'team of monkeys' who produce random trial maps by scattering quanta of radio intensity across the cells of the sky. This is done without regard to observations, and produces a large pile containing, many times over, all positive trial maps of the sky. We take their Fourier transforms, and compare these with the data. Almost all of the trial maps will be inconsistent with the data despite the observational errors, and these we reject. Those which are left, and are compatible with the given data and the errors, are

*Permanent address: Department of Physics, The University, Southampton, UK.

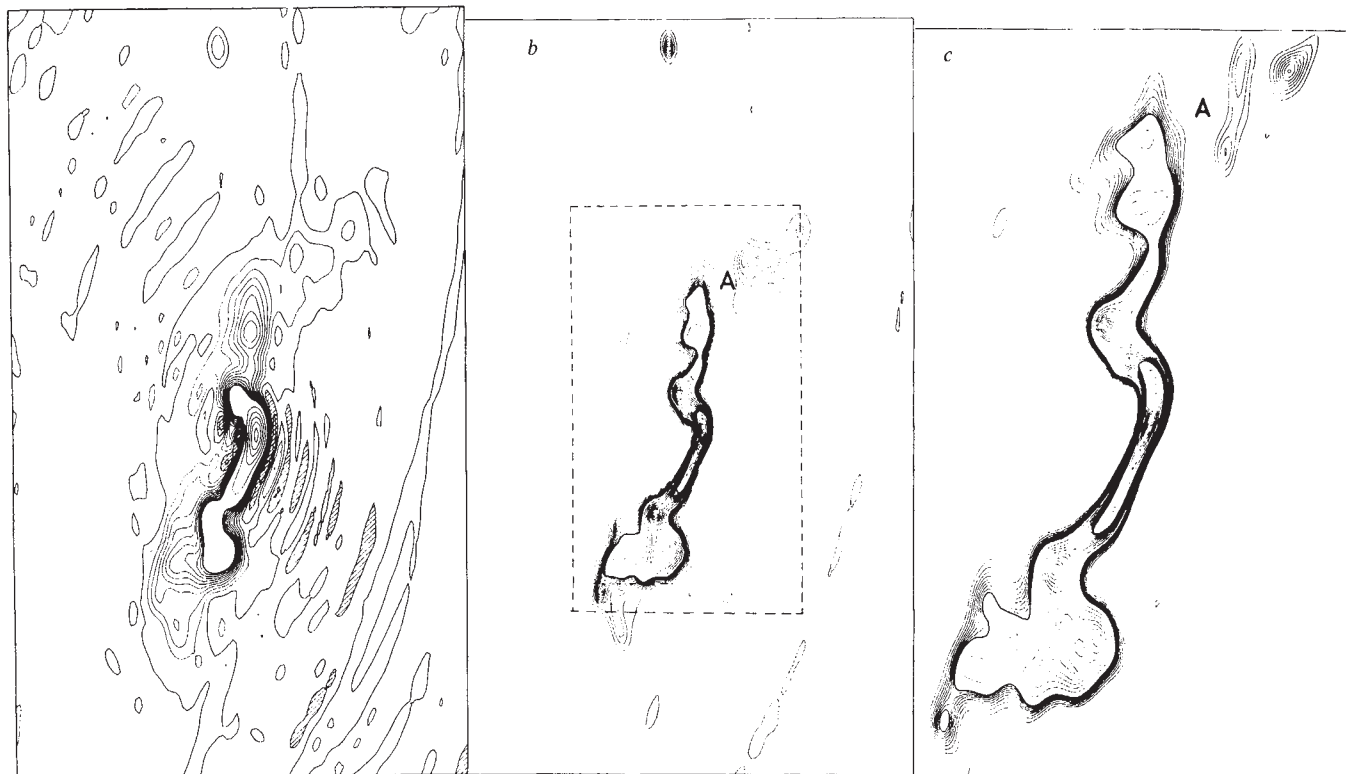


Fig. 1 *a*, Conventional map of the radio galaxy 3C31. The contour interval is increased by a factor of 10 after the first 10 contours. Negative regions are shown hatched. *b*, Maximum-entropy map of 3C31, on the same scale as Fig. 1*a*. The contour interval is increased by a factor of 10 after every set of 10 contours, so that there are 750 of the lowest contours to the peak. The box shows the region covered by Fig. 1*c*. *c*, The central region of 3C31, analysed by the maximum entropy algorithm. The contour levels are the same as in Fig. 1*b*.

sorted into piles, different maps being placed in different piles. Because we have quantised the space and intensity levels, there are only a finite number of such piles. We continue this process until each conceivable, compatible map has turned up many times, and then select the maps in the largest pile as the most likely representation of the sky.

The map most frequently produced by the monkeys has a uniform intensity. This procedure, therefore, produces the most uniform map consistent with the data. Indeed, if the noise in the data is larger than any real signal from the sky, then the uniform map will be consistent with the data and is thus the maximum-entropy solution. Conversely, if there is a feature on a maximum-entropy map, and the errors have been correctly estimated, then there must be evidence for it in the data. In this sense the method is safe. However, it does not follow that features on a maximum-entropy map must correspond to real features in the sky. We show later an example where the true noise level is much greater than the value used in the analysis, and features arise from the interpretation of this noise as real signals. Similarly, features can arise from an ambiguity in incomplete data over two or more possible interpretations. In such cases the method selects the smoothest map, which is not necessarily correct.

For simplicity we write the equations for data in one dimension, the extension to two or more dimensions being immediate. Let m_j denote the intensity at point j of a test map and let M_k be the Fourier transform of m_j . Suppose that we have measurements of M_k , denoted by E_k , only on a subset of k -values, which we call A . There is no restriction on the domain A , which can be any set of points in k -space.

There is some dispute about the correct definition of the entropy. We believe that the arguments of Frieden⁵ are correct for our problem. He concludes that the most probable map is that which maximises $-\sum_j m_j \log m_j$. This formula expresses the configurational entropy of a map; this is not the same as the thermodynamic entropy of a beam of photons, nor is it the same as the information-theoretical entropy loss on passing a signal through a filter. An alternative expression, $\sum_j \log m_j$ (ref. 6), has

been widely used in maximum-entropy spectral analysis and has also been applied in radioastronomy. We find the arguments leading to this expression less compelling than those of Frieden for the radioastronomical problem. We shall discuss the foundations of the method further elsewhere.

The unconstrained maximisation of $-\sum_j m_j \log m_j$ results in all the m_j being equal to e^{-1} , and maximising it subject to the constraint that the M_k are consistent with the data produces our solution. It is this approach which differs from that of some others^{5,7}. We do not attempt to fit the data exactly, because this introduces structure into the map arising solely from noise in the data. Exact fitting also implies the existence of numerous separate constraints, resulting mathematically in an unwieldy proliferation of Lagrange multipliers and preventing calculation of the solution in all but the simplest cases.

We assume instead that the data E_k have gaussian errors, with individual standard deviations σ_k . The statistic $\sum_{k \in A} |M_k - E_k|^2 / \sigma_k^2$ then possesses a χ^2 distribution. The sum is taken over the set of points A at which there are observations. A large value of χ^2 indicates that the map is a poor fit to the data, while a very small value arises when the map fits the data too closely and includes features due solely to experimental error. The expected value of χ^2 is equal to the number of data values, and we use Lagrange's method of undetermined multipliers to constrain χ^2 to equal this while we maximise the entropy. Maximising

$$Q(\lambda) = -\sum_j m_j \log m_j - \frac{\lambda}{2} \sum_{k \in A} |M_k - E_k|^2 / \sigma_k^2$$

with respect to m_j gives

$$m_j = \exp \left\{ -1 + \lambda \sum_{k \in A} (E_k - M_k) / \sigma_k^2 \exp(2\pi i j k / N) \right\} \quad (1)$$

The multiplier λ must be positive; the apparently arbitrary

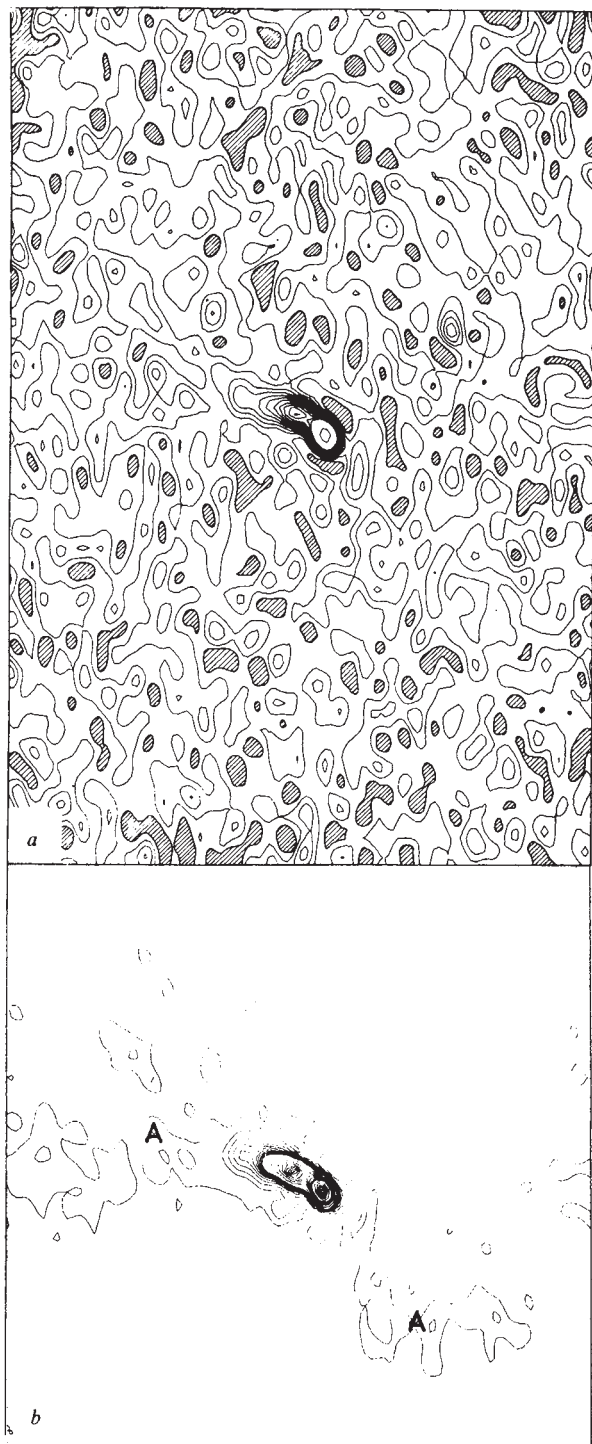


Fig. 2 *a*, Conventional map of 3C66. The contours are confused by fluctuations in the background, which are the result of receiver noise. Negative regions are shown hatched. *b*, Maximum-entropy map of 3C66. Regions of low surface brightness are now visible (A).

constant e^{-1} is related to the problem of the scaling of the data. This problem can be resolved by a more careful investigation of the foundations of the method.

Given E_k , σ_k and λ , this equation determines the intensities m_j on the map. It has been solved numerically by an iterative procedure, starting with the uniform map. Each new iterate is obtained by evaluating the right hand side of equation (1) with M_k equal to the Fourier transform of the previous iterate. We must average successive iterates to obtain convergence. The solution for a map of 128×128 grid points takes 3 min on the IBM 370/165

computer of the University Computer Laboratory, Cambridge. We can show that a solution of equation (1) must exist for any $\lambda > 0$, that it is unique and that it makes the quantity $Q(\lambda)$ a global maximum.

The correct value of λ makes χ^2 equal to the number of observations. During our iterative solution, λ is automatically increased until this is achieved by making it dependent on the current value of χ^2 . In practice the solution is almost unchanged for a wide range of λ .

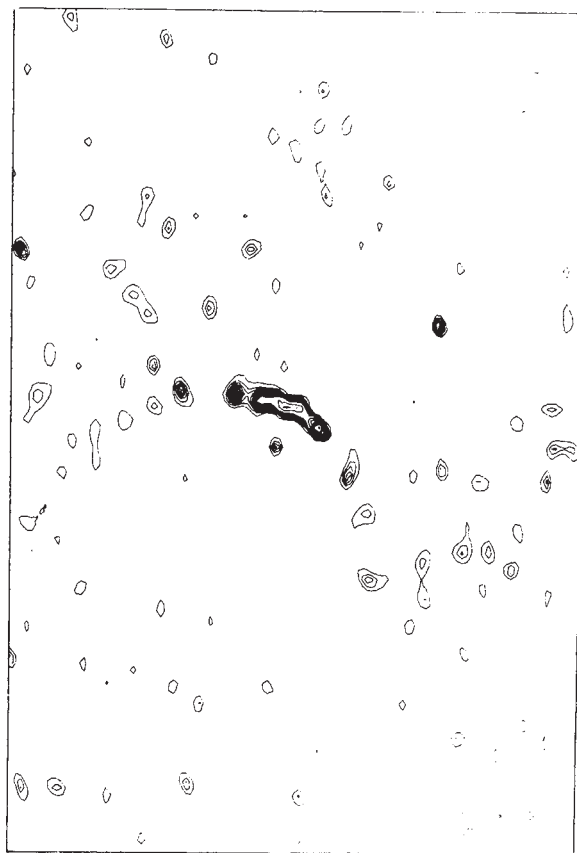
Applications using radioastronomical data

We have made extensive tests with simulated data. In every case the maximum-entropy reconstruction has been very similar to the original and we have found no cases where a misleading conclusion could have been drawn. We therefore present here some results using real observations made by the Cambridge radio telescopes. These data are incomplete because of the finite size of the telescope, because the data may be spoiled by electrical interference, or because the expense of further data collection outweighs the improvement to the final map of the sky. Errors in the data vary from sample to sample because of differing integration times and atmospheric conditions.

Figure 1*a* is a conventional map of the radio galaxy 3C31, observed at 1,407 MHz by the One-Mile Telescope⁸. The dynamic range is about 50:1. The shaded areas indicate negative regions, which arise principally as sidelobes of the main peak. Note the change in contour interval by a factor of 10 after the first 10 contours.

Figure 1*b* shows the maximum-entropy map produced with the same data. All negative sidelobes have been removed and positive ones greatly reduced, because these sidelobes arise from missing Fourier components and can be removed without affecting the fit to the data. The central peak is sharper and the noise background

Fig. 3 An example of the mis-use of the maximum-entropy method. The spurious features, which are not present in Fig. 2*b*, result from an over-optimistic estimate of the noise level in the data.



has been reduced, so that the apparent dynamic range is 1,000:1. An interesting feature is the valley A in Fig. 1b, which is barely visible in Fig. 1a. As our algorithm generates the flattest map consistent with the data we conclude that there is little evidence in the observations for a bridge connecting these two parts of the radio source. There may be one, but we need better data to see it.

A more detailed maximum-entropy map of the central part of this source is shown as Fig. 1c. To produce this we assume that our data come from the central part of a much larger radio telescope and that the Fourier components corresponding to the outer parts of the aperture are unknown. The algorithm now allows us to assign values to these unobserved Fourier components, but only in a way which makes the final map flatter. We see a further striking reduction in confusion caused by sidelobes, and a reduction in width of the main ridge of the source by a factor of about 2. This increase of resolution is attainable because, in Fig. 1a, the Fourier components corresponding to large spacings

'super-resolution'. We include this picture to emphasise the dangers inherent in over-interpreting observational data. For the maximum-entropy method these dangers can be minimised by plotting a map of $|E_k - M_k|^2/\sigma_k^2$. Corrupted data then become obvious through their unacceptably large contributions to χ^2 .

Maximum-entropy solutions without phase information

We have so far assumed that the observations E_k are complex numbers, and that their real and imaginary parts have equal and independent gaussian errors. This is unrealistic in many important radioastronomical applications. For example, at high frequencies the main source of uncertainty is the troposphere, which introduces unknown phase errors into the Fourier components without affecting their amplitudes. In extreme cases, such as in very long baseline radio interferometry or in X-ray crystallography, the phases are entirely unknown.

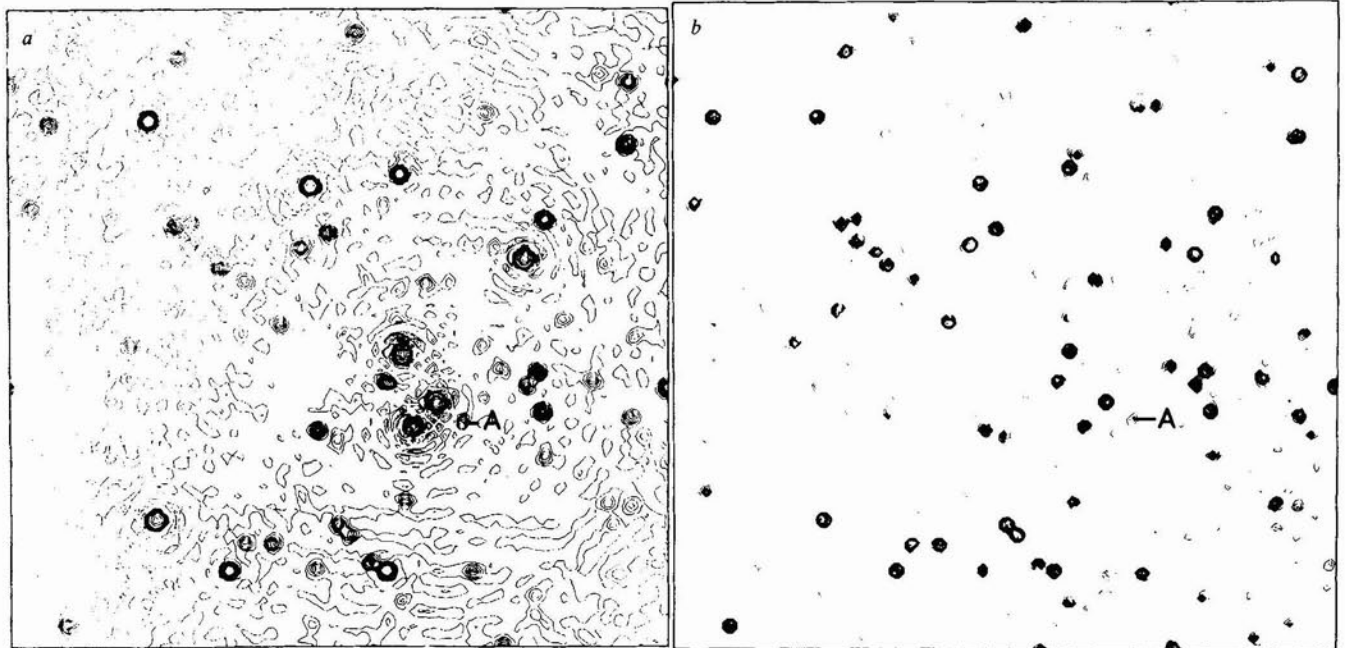


Fig. 4 a, Part of the 5C7 survey of radio sources. b, Phaseless maximum-entropy map of 5C7.

have been weighted down to one-third of their measured values to improve the sidelobe level, despite the consequent loss of resolution. In the case of 3C31, the extra details revealed by the algorithm are well confirmed by observations with higher resolving power⁸. A modest amount of 'super-resolution' is possible with this method if the signal-to-noise ratio is sufficiently high but any further increase in the size of the apparent aperture does not result in greater resolution. Because we consider noise in the data, we cannot get the implausibly high resolution that is suggested by some other methods of maximum-entropy spectral analysis⁷.

Figure 2a shows the central region of the radio galaxy 3C66, observed at 5GHz by the One-Mile Telescope⁹. These data are much more noisy than those for 3C31 and the maximum-entropy map (Fig. 2b) shows only a very small increase in resolution. However, the map now shows clear evidence for more diffuse regions of lower surface brightness (marked A), as these are no longer confused by noise. The reality of these features is confirmed by observations at lower frequencies⁹, where receiver noise is less of a problem.

Figure 3 shows a typical result if the experimental errors are seriously underestimated. To produce this unfortunate result, we assumed that each σ_k was five times smaller than that used to produce Fig. 2b. Noise in the data has now been interpreted as true signal and appears on the map. There is also some spurious

Under these circumstances, we use $\sum_{k \in A} (|M_k| - A_k)^2/\sigma_k^2$ as a measure of the misfit to the data, where A_k denotes the observed amplitudes. Although this quantity does not have a χ^2 distribution, the number of observations is usually so large that we may apply the central limit theorem. The distribution is therefore approximately gaussian and its expectation value is easily estimated.

Maximising the entropy subject to the constraint that the misfit is equal to its expectation value yields

$$m_j = \exp\{-1 + \lambda \sum_{k \in A} (A_k \exp(i \arg M_k) - M_k)/\sigma_k^2 \exp(2\pi i j k/N)\}$$

which may be solved by iteration in the same way as before. This time the solution is not unique; as well as the trivial cases of arbitrary translation or inversion of the solution, we have a technical counter-example (due to Dr J. Skilling) suggesting that the entropy may sometimes have more than one local maximum. We have, however, encountered no such ambiguity in our tests with realistic radioastronomical data.

Figure 4a shows part of the radio survey 5C7 (ref. 10), made with the One-Mile Telescope at 1,407 MHz. This map has been produced in the conventional manner, using the observed amplitudes and phases. The maximum-entropy map (Fig. 4b) has been

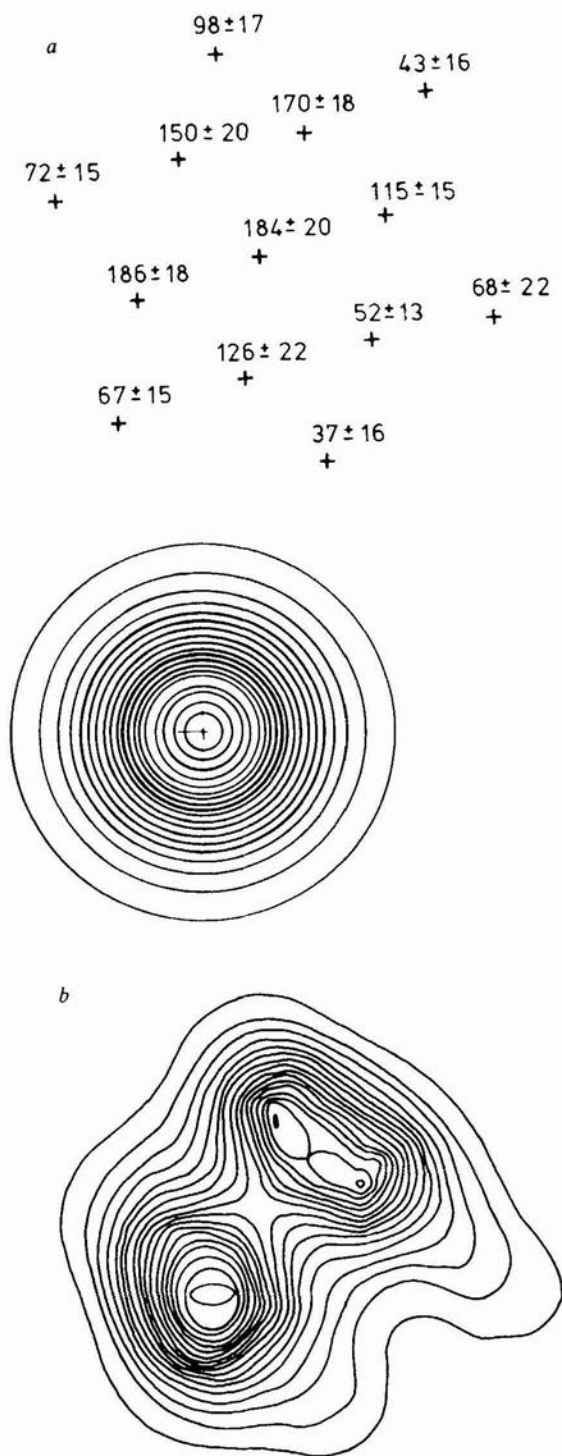


Fig. 5 *a*, X-ray counting rates for various positions near the supernova remnant Cassiopeia A. Also shown are the experimental errors, and a contour map of the point-spread function. *b*, Maximum-entropy X-ray map of Cassiopeia A on the same scale as Fig. 5*a*.

made utilising only the amplitude information. The phases were in fact given to the program, but only to start the iteration near the correct solution and to encourage it to produce its solution in the same orientation and position as Fig. 4*a*. Test cases show that the algorithm produces the correct solution even if no phase information whatever is given, although convergence is slower. All the sources visible on the conventional map have been reproduced, and most have the correct positions and flux densities. A few of the weak sources (< 1% of the brightest) have a slightly lower flux

density on the maximum-entropy map; an example is marked A. Presumably, evidence for the presence of this weak source cannot be extracted from these very incomplete data in the absence of phase information. The same radio survey has also been analysed by Baldwin and Warner¹¹, using a method similar to those used in X-ray crystallography. Both reconstructions closely resemble the true map, but ours has the anticipated features of good resolution and large dynamic range.

Further applications

The combination of maximum-entropy analysis with a proper treatment of experimental errors thus promises to be a powerful, yet safe, method for estimating the Fourier transform of incomplete and noisy data. For radioastronomical data, the scheme offers considerable advantages over other reconstruction techniques, for example, the CLEAN algorithm¹². Because our algorithm produces a well-defined map even if the phases are entirely unknown, it is also ideal for the reconstruction of images produced by very long baseline radio interferometry, X-ray crystallography and speckle-pattern interferometry. Full technical details of our algorithm for the case of Fourier data will be given elsewhere.

The method is very easily generalised to deal with other types of data, not involving a Fourier transform. A particularly important example is the case where the data collected ($E_k \pm \sigma_k$) are the convolution of the required map with a known point-spread function or beamshape. The extension of the maximum-entropy method to this case is quite straightforward and yields

$$m_j = \exp\{-1 + \lambda \sum_{k \in A} B_{k-j} (E_k - g_k) / \sigma_k^2\}$$

where $g_k = \sum_j m_j B_{k-j}$ and B_j is the point-spread function. The maps are again smooth and give resolution limited by the signal-to-noise ratio. In this sense the solutions are far superior to those produced by the commonly used 'algebraic reconstruction technique'¹³. The maximum entropy-solution is again unique.

Immediate applications of this deconvolution technique exist in many branches of physics. As an example, we show an X-ray map of the supernova remnant Cassiopeia A, observed by the Copernicus satellite¹⁴ and previously reconstructed by the ART algorithm¹⁵. The map is based on only 13 data points, which are shown in Fig. 5*a*, together with the point-spread function. The maximum-entropy map is shown as Fig. 5*b*, and resolves the supernova remnant into two major regions of X-ray emission. This reconstruction in fact resembles that produced by the ART algorithm¹⁵, but is virtually identical to the radio map of the supernova at the same resolution.

In this simple example only 64×64 grid points are needed to show the full details of the maximum-entropy solution. However, the computer time required implies that there is no difficulty in processing pictures of up to 256×256 grid points and therefore that that digital processing of photographs by this algorithm is realistic with a fast computer.

We thank the Radio Astronomy Group for helpful comments, also Dr J. Skilling for proof of the existence and uniqueness of maximum-entropy solutions.

Received 2 November 1977; accepted 12 January 1978.

1. Parker, R. L. *A. Rev. Earth Planet. Sci.* **5**, 35 (1977).
2. Oldenburg, D. W. *Geophys. J. R. astr. Soc.* **44**, 413 (1976).
3. Jaynes, E. T. *Phys. Rev.* **106**, 620 (1957).
4. Jaynes, E. T. *Phys. Rev.* **108**, 171 (1957).
5. Frieden, B. R. *J. opt. Soc. Am.* **62**, 511 (1972).
6. Abels, J. G. *Astr. Astrophys. Suppl.* **15**, 383 (1974).
7. Burg, J. P. thesis, Stanford Univ. (1975).
8. Burch, S. F. *Mon. Not. R. astr. Soc.* **181**, 599 (1977).
9. Northover, K. J. F. *Mon. Not. R. astr. Soc.* **165**, 369 (1973).
10. Pearson, T. J. & Kus, A. J. *Mon. Not. R. astr. Soc.* **182**, 273 (1978).
11. Baldwin, J. E. & Warner, P. J. *Mon. Not. R. astr. Soc.* **182**, 411 (1978).
12. Höghom, J. A. *Astr. Astrophys. Suppl.* **15**, 417 (1974).
13. Stevens, J. C. & Garmire, G. P. *Astrophys. J.* **180**, L19 (1973).
14. Fabian, A. C., Zarnecki, J. C. & Culhane, J. L. *Nature phys. Sci.* **242**, 18 (1973).
15. Charles, P. A., Culhane, J. L. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **178**, 307 (1977).

Morphology of martian rampart craters

Peter J. Mouginis-Mark*

Lunar and Planetary Unit, University of Lancaster, Lancaster, UK

Preliminary results from a morphological analysis of the lobate flows around martian rampart craters are presented. The areal dispersal of these deposits displays a proportionally greater maximum range for the lobe as the size of the parent crater increases. It is proposed that lobe formation is not entirely associated with ballistic deposition of the crater's ejecta because the proximal ends of the lobes are found closer to the primary crater rim than to the region of secondary cratering associated with either lunar or mercurian craters. Lobe surface area is shown to be proportional to crater diameter and an estimate of 30–60 m is deduced for the lobe thickness for craters between 10 and 35 km diameter.

THE Viking missions to Mars in 1976 provided high resolution photography of many new types of landforms. Of particular interest is a group of craters exhibiting lobate deposits around their rims instead of the normal 'blocky' ejecta blankets observed for lunar and mercurian impact craters¹. Termed a 'rampart' crater, because of the sharp edge between the distal end of the flow and the surrounding terrain, this type of crater was first observed from Mariner 9 photography². Its unusual character has been attributed either to wind deflation of a ballistically emplaced ejecta blanket^{2,3}, or the radial movement of a debris flow from the parent crater as a late stage process in the crater-forming event^{1,4,5}. A prerequisite in all the debris flow models has been the existence of a layer of ground water or permafrost which lubricates the crater's ejecta after its initial emplacement. Retention of target volatiles, air-layer cushioning as a result of the entrainment of the tenuous martian atmosphere (comparable to the transport mechanism of certain terrestrial landslides⁶) or the incorporation of volatiles during the formation of an impact-generated base surge⁷, have all been proposed as possible mechanisms by which the ejecta might increase its mobility in comparison with impact events on the Moon and Mercury^{1,4,5}.

Possible terrestrial equivalents to these lobate flows have been discussed elsewhere⁸. This paper describes several novel aspects of the morphology of this type of crater. Sixty-three craters ranging from 2 to 38 km in diameter have been included in the sample. The mean diameter (the average of the north-south, east-west, north-west-south-east and north-east-south-west diameters) the maximum and minimum range of the lobes from the rim of the parent crater, and, where suitable Viking imagery was available, the surface area of the deposit, were measured.

Range of the lobate flows

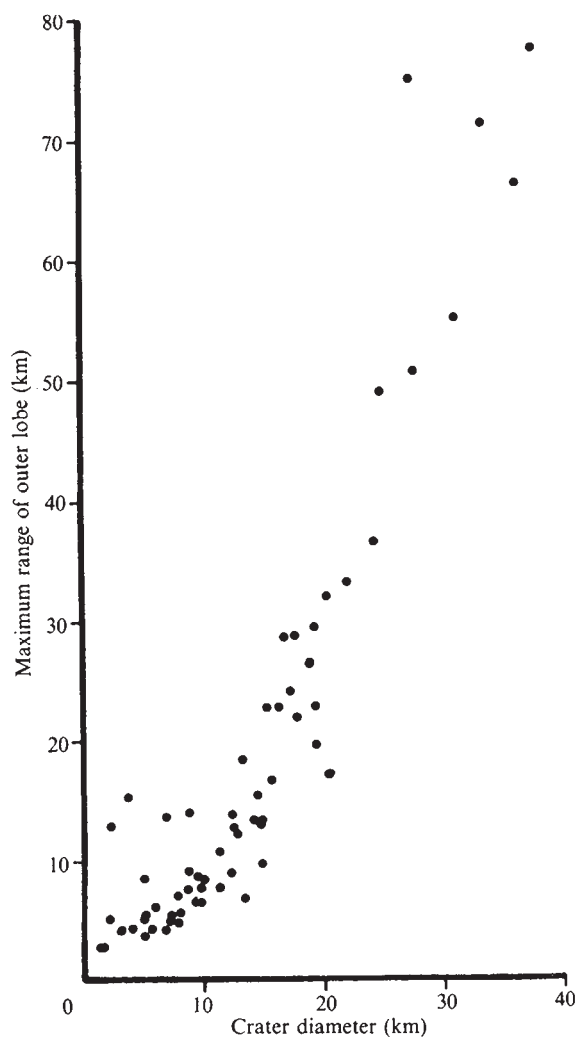
The mean radial extent of the continuous ejecta from rampart craters, expressed in units of the diameter of the parent crater, has already been described⁹. Prominent lobes were found to extend to three times the range from the rim of the parent crater than the continuous ejecta of mercurian craters, which are considered to be comparable to the martian craters had they too formed on an airless planet⁹. More detailed information pertinent to the mode of formation of the rampart crater lobes can be derived from the distal range of the deposits because this property

relates directly to the mobility of the flow at the time of formation.

Figure 1 shows the maximum range of each lobate deposit as a function of the diameter of the parent crater. For primary craters >10 km diameter, an almost linear correlation exists between maximum lobe range and crater diameter, whilst for smaller primaries the maximum lobe range is approximately 15 km from the rim of the parent crater whatever the diameter. Expressing the maximum range of each lobe in units of the primary crater's diameter (Fig. 2) reveals that for craters larger than 10 km diameter, the maximum range of the lobe from the rim crest of the parent crater can be approximated by: $R/D = D/10^{\pm}$ where R is the maximum range of the lobe from the crater rim and D is the diameter of the crater, both in km.

A similar, but numerically different, relationship has also been reported between the volume of material in, and the distance travelled by, terrestrial landslides¹⁰, whilst an in-

Fig. 1 Maximum range of rampart crater lobes as a function of primary crater diameter.



*Present address: 26913 Cook Road, Olmsted Twp., Ohio 44138.

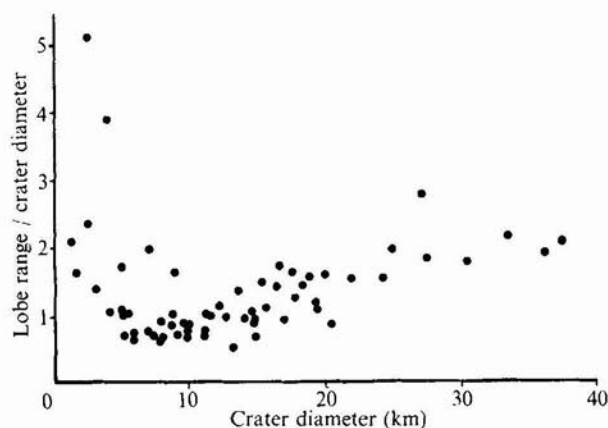


Fig. 2 Maximum range of rampart lobes as a function of primary crater diameter: normalised to the diameter of parent crater. Craters larger than 10 km diameter illustrate a maximum lobe range which approximates to: $L/D = (D/10)^{1/2}$ where L is the lobe range and D the crater diameter.

crease in the mobility of pyroclastic flows may also exist as the volume and 'fall height' of an ignimbrite increases¹¹. It seems, therefore, that the increased mobility of the lobes around larger rampart craters is a product of the volume of the material excavated from the crater cavity, rather than being a physical manifestation of a higher flow moisture content.

Craters smaller than 10 km diameter conform to a different relationship: the maximum travel distance of a lobe is limited to no more than 15 km from the rim of the parent crater, but the range is independent of crater diameter. For this size of crater, it is possible that lobe dispersion may be controlled more by the local topography and the structural uplift of the crater rim. Because the lobate flows from these smaller craters are morphologically similar to the flows around large rampart craters, it is believed that they are both part of the same population, although the precise mechanism of lobe formation for small craters is not well understood.

Secondary impact cratering has been suggested as a possible mechanism for the initiation of lobe movement¹. When large secondary blocks of ejecta impact the surrounding terrain, large quantities of ground moisture may be incorporated in a martian 'muddy base surge' that gains some of its mobility from the moisture-rich tertiary ejecta. It is suggested here and discussed in more detail elsewhere⁸, however, that secondary cratering is not the process by which the mobility of the flows is controlled. Stratification within the lobes around many of the larger rampart craters to form two or more near continuous lobes implies more than one period of lobe formation, because the inner lobes appear to over-ride the earlier, more extensive, outer lobes. Size sorting of in-flight ejecta, by atmospheric drag, has been proposed as a possible process by which two or more discrete periods of secondary cratering may occur on Mars¹⁴. In this model, however, ejecta blocks were considered to travel independently from the point of excavation to the final point of impact, and this bears little relation to the most likely process of ejecta emplacement on Mars because it assumes that the atmosphere is able to move around each block of ejecta. Due to the very high ejecta cloud density/atmospheric density ratio, ejected blocks cannot interact individually with the atmosphere, and little atmospheric deceleration of an ejecta cloud should occur⁸, so that the deposition of martian ejecta may be similar to the lunar case⁷.

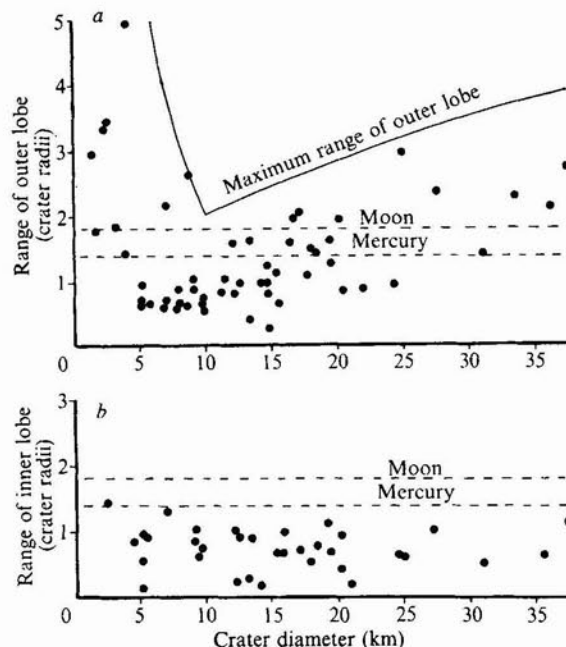
The closest approach to the rim of all the rampart crater lobes (and, where they exist, the inner lobes as well) are shown in Fig. 3 as a function of crater diameter. Also

depicted are the distances at which maximum secondary cratering around lunar and mercurian craters is observed^{5,9}. It is clear that many of the outer rampart lobes, and all of the inner lobes, originate much closer to the primary crater than the region where maximum secondary cratering occurs. If secondary cratering were the dominant process by which lobes were initiated, the distribution presented in Fig. 3 should not occur until at least 1.4 crater radii from the primary crater because this is the region where the process is predicted to be most influential^{5,9}. It is, therefore, unlikely that secondary cratering plays a major part in the process of rampart crater lobe formation.

Surface area of rampart lobes

Photographs of only 14 craters were available for the measurement of the surface area of the rampart lobes. The sample is small because it was considered necessary to have at least 270° azimuthal coverage of the entire ejecta blanket on orthographic, or near vertical, high resolution Viking imagery. Lobe surface area was calculated by tracing the lobe boundary onto a clear acetate overlay and placing this overlay on top of a sheet of graph paper and counting the underlying squares (each of which were of known relative size with respect to the crater being measured). Craters with less than 360° azimuthal coverage were given a scaled lobe surface area. In all cases, only the surface area of the inner lobe (or the single lobe for the smaller craters) was measured, except for two craters where superior photography permitted the outer lobe (Fig. 4a) as well as the inner lobe to be included. When the logarithm of the surface area of the lobes is plotted against the logarithm of the crater diameter (Fig. 4) a surprising correlation is revealed. With the exception of the craters *a*, *b* or *c* in Fig. 4, the relationship between the two variables is almost linear: when the crater diameter is plotted as the abscissa the slope of the straight line has a value of 3.06. The exclusion of crater *b* is justifiable because it lies close to the distal end of a large channel system and hence may represent a crater formed in an abnormally moist environment. Crater *c* can also be

Fig. 3 Minimum range of outer (*a*) and inner (*b*) lobes. Also shown are the locations of maximum secondary cratering for the Moon and Mercury^{5,9}, and the maximum range of the outer lobes given in Fig. 2. It is demonstrated that lobe initiation occurs much closer to the rim crest of the primary crater than the area of maximum secondary cratering, and so lobe formation is unlikely to be dominated by the secondary cratering process.



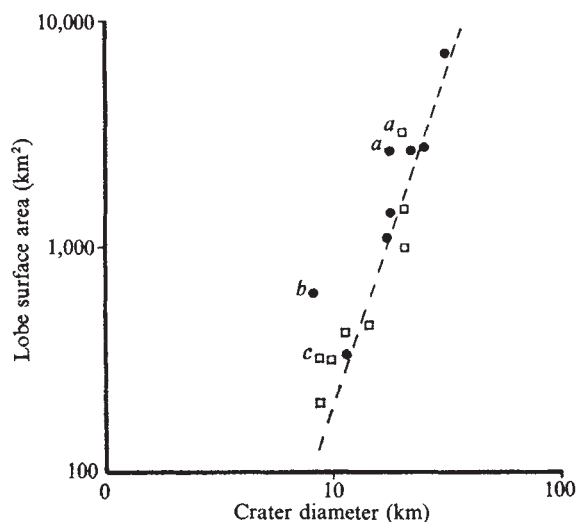


Fig. 4 Rampart lobe surface area as a function of crater diameter. Data derived from two sources: orthographic projections of Viking imagery (●) and near vertical rectilinear prints (□). *a*, Two measurements of the outer lobes of craters, all other values are for the inner lobes of twin-lobed craters and the single lobe for smaller (<15 km) craters; *b* has an abnormally high (surface area/crater diameter) ratio, due to the proximity of a delta system. *c* is a doublet crater with both craters contributing to the total lobe surface area. Craters *a*, *b* and *c* have been omitted in determining the mean thickness of the lobate deposits. The dashed line corresponds to the function: lobe thickness $\propto$ crater radius $^{-0.06}$.

excluded because there is a second, smaller primary crater on the lobate deposit of the large primary which has also contributed to the surface area of the entire deposit.

From Fig. 4, we find that the area of the lobe is given by the function $A = 0.152 D^{3.06}$, where A is the surface area of the lobe in km^2 and D is the diameter of the crater in km. Because fresh impact craters display a close association between depth and diameter^{13,14}, it follows that there is some correlation between crater diameter and the volume of the excavated crater cavity (and hence the ejected material). If this volume is given as v , and the crater radius as r , then: $v \propto r^3$. Assuming that the rampart crater lobe forms a layer of mean thickness t and surface area A , it follows that: $A t \propto r^3$. From Fig. 4, it is shown that $A \propto r^{3.06}$ and so $t \propto r^{-0.06}$.

It seems, therefore, that the mean thickness of the rampart lobes is nearly independent of crater size, and that some mechanism apart from the cratering event determines the thickness of the lobes. Figure 5 represents the volume of a rampart crater lobe deposit for different estimates of the mean lobe thickness. Also shown are five pairs of data points for five morphologically fresh mercurian craters. The dimensions of these craters were computed by photoclinometric techniques⁸ similar to those used by Hapke *et al.*¹⁵ and display similar depth/diameter ratios to those of fresh martian craters¹⁶. The volume of the excavated crater cavity and the crater rim were calculated for these mercurian craters from the measured profiles using the equations of Pike¹⁷. It is assumed that mercurian craters are acceptable comparisons with martian craters, because the gravitational fields of the two planets are similar^{5,9}. The photoclinometric technique could not be applied to the martian craters due to random photometric variations in the Viking imagery which were induced by the martian atmosphere.

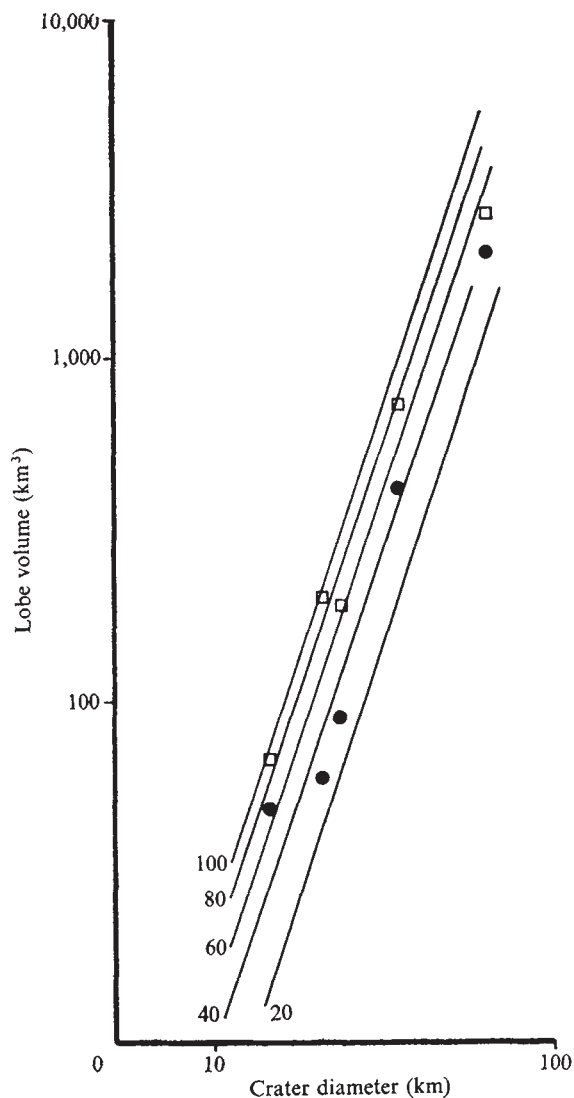
Figure 5 demonstrates that if the 'missing volume' between the volume of the excavated crater cavity and the volume of the crater rim (which represents both the structural uplift and deposited ejecta) is dispersed uniformly around the rampart crater, a mean thickness for the rampart lobes

between 30 and 60 m is likely. All five values for the missing volume of the mercurian craters lie within this region. Such a value for the lobe thickness compares favourably with other planetary flow features. The Blackhawk Landslide in Southern California is ~15 m thick⁶ and the landslide feature on the rim of the lunar crater Tsiolkovsky is estimated to be ~100 m thick¹⁸.

Conclusions

Rampart craters represent a new category of martian craters. It is assumed here that the main reason for their unusual morphology is the presence of a layer of ground moisture (either a martian aquifer or a layer of permafrost) at the time of crater formation. The increase in the maximum range of the lobes with increasing crater diameter reflects a ground flow process which may have been similar to a mudflow or landslide¹⁰. The uniformity of the mean lobe thickness for a range of crater diameters precludes a simple ballistic deposition origin for the lobe, and this model is also rejected on the grounds that multiple-lobe initiation is observed to occur very close to the primary crater's rim. Future analysis of rampart craters should re-

Fig. 5 Lobe volume lines (in m for the estimate of the thickness of the lobe) deduced from the relationship shown in Fig. 4. Also shown are the photoclinometrically derived¹⁵ estimates for the crater cavity volume (□) and cavity volume minus rim volume (●) of five morphologically fresh impact craters on Mercury. A mean lobe thickness for all sizes of rampart craters is inferred to be ~30–60 m, as this corresponds to the missing volume of the ejecta from the mercurian craters.



flect the wide variety of possible modes of formation for these features. Results relating to the possibility that these lobes are produced by heating from a hot ejecta blanket emplaced on a layer of permafrost, mass-movement under martian periglacial conditions and the influence of various ground moisture contents on the size of the lobes will be presented in the near future.

This research was carried out during the tenure of a NERC studentship. I thank Drs Wilson and Fielder for their advice and encouragement over the last three years, and the members of the Lunar and Planetary Unit at Lancaster for their helpful suggestions. The work on the Viking data forms part of a joint investigation with the University of London Observatory at Mill Hill.

Received 16 December 1977; accepted 1 February 1978.

1. Carr, M. H. *et al. Science* **193**, 766–776 (1976).
2. McCauley, J. F. *J. geophys. Res.* **78**, 4123–4138 (1973).
3. Arvidson, R. E. *et al. Icarus* **27**, 503–516 (1976).
4. Head, J. W. & Roth, R. in *Symp. planet. Cratering*, Flagstaff Arizona (1976).
5. Carr, M. H. *et al. J. geophys. Res.* **82**, 4055–4065 (1977).
6. Shreve, R. L. *Geol. Soc. Am. Sp. Pap. No.* 108 (1968).
7. Oberbeck, V. R. *Rev. geophys. Space Phys.* **13**, 337–362 (1975).
8. Mouginiis-Mark, P. J. thesis, Univ. Lancaster (1977).
9. Gault, D. E., Guest, J. E., Murray, J. B., Dzurisin, D. & Malin, M. L. *J. geophys. Res.* **80**, 2444–2460 (1975).
10. Hsu, K. J. *Geol. Soc. Am. Bull.* **86**, 129–140 (1975).
11. Sparks, R. S. J. *Sedimentology* **23**, 147–188 (1976).
12. Schultz, P. H. & Gault, D. E. *Proc. 8th Lunar Sci. Conf.*, Houston, Texas (1977).
13. Pike, R. J. *The Moon* **15**, 463–477 (1976).
14. Pike, R. J. *Proc. 7th Lunar Sci. Conf. Abstr.* 700–702 (1976).
15. Hapke, B., Danielson, E., Klassen, K. & Wilson, L. J. *J. geophys. Res.* **80**, 2431–2443 (1975).
16. Burt, J., Veverka, J. & Cook, K. *Icarus* **29**, 83–90 (1976).
17. Pike, R. J. *J. geophys. Res.* **72**, 2099–2105 (1967).
18. Guest, J. E. in *Geology and Physics of the Moon* (ed. Fielder, G.) (Elsevier, Amsterdam, 1971).

Membrane leakiness after viral infection and a new approach to the development of antiviral agents

Luis Carrasco*

Translation Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2, UK

Viral development induces changes in the permeability properties of the plasma membrane of the host cell. Here it is shown that, because of this leakiness, inhibitors of protein synthesis normally impermeable to uninfected cells are able to enter infected cells and thereby specifically inhibit viral protein synthesis.

AFTER a cytolytic virus attaches to specific membrane receptors, its nucleic acid penetrates into the cell and starts to express the viral genome, thereby interfering with cellular functions^{1–4}. However, the exact mechanisms by which the virus is able to take over the cellular metabolism remain unknown. We have recently proposed that one factor in the strategy for viral take-over is the modification of the plasma membrane⁴. Based on the fact that the optimal ionic conditions for cellular and viral protein synthesis *in vitro* are different⁵, we proposed that the preferential translation of viral mRNAs occurs physiologically in a cytoplasm in which there is a gradual increase in monovalent ion concentration. If this is so, then the shut-off of host cell functions, the bulk of the translation of viral mRNAs with higher monovalent ion requirement for *in vitro* translation, and the changes in the ionic concentration within the cell, should be concomitant processes. The membrane-leakage model⁴ supposes that the modification of the membrane is caused by a viral component. Such a change not only involves a gradual redistribution of ions but also one of small molecules between the surrounding medium of the cell, the cellular organelles and the cytoplasm itself, and should affect in parallel a spectrum of functions dependent on membrane integrity.

In support of this model, animal viruses^{6–13} as well as bacteriophages^{14–17} alter the membranes of their host both structurally and functionally, and these alterations lead to changes in the distribution of ions and metabolites between the cell interior and the surrounding medium^{5,10,18–20}. If viral development involves membrane leakiness, we need

to know (1) what mechanisms the virus uses to induce such a modification in the membrane, (2) at what time after infection the modification of the membrane occurs, (3) whether there is a general strategy followed by all viruses, and (4) whether specific antiviral agents could be designed to make use of the modification of membranes caused by viral development.

Using impermeable inhibitors to test for membrane permeability

It could be argued that the membrane of a virus-infected cell only becomes leaky after cell death. In fact, the most commonly used assay for cell viability is the stainability of cells by a dye which is normally excluded from cells. Such a test has been recently questioned¹³, because although SV40-infected CV1 cells can be stained by Trypan blue during the late phase of the infection and therefore might be considered dead, the cells are still functioning in many other respects. Also, in chick embryo fibroblasts infected with Newcastle disease virus, molecules such as sucrose can enter the cell late after infection⁶. Similarly, ions and small molecules like ATP start to leak out from bacteriophage¹⁵ and picornavirus-infected¹⁹ cells late in infection. However, from these experiments we cannot distinguish whether only a percentage of the infected cell population has a leaky membrane and that these cells have already stopped making viral proteins and producing virus, or whether all cells gradually become leaky as they reach a specific step in the late phase of infection.

A suitable test to distinguish between these two possibilities would be to use a low molecular weight inhibitor of protein synthesis to which cells are not normally permeable. As nucleotides and ionised molecules in general do not enter normal cells²¹, we chose as an impermeable inhibitor of protein synthesis the GTP analogue GppCH₂p. With this compound we tested whether or not an infected cell which is still synthesising viral proteins has a leaky membrane as shown in Figs 1 and 2, the addition of this nucleotide analogue to the medium of mock-infected 3T6 cells had no effect on cellular protein synthesis, whereas it markedly inhibited translation in EMC-infected 3T6 at a time when viral protein synthesis was maximal. Therefore, assuming that GppCH₂p does not inhibit protein

*Present address: Insto, Bioquímica de Macromoléculas, Centro de Biología Molecular, Universidad Autónoma de Madrid, Facultad de Ciencias, Canto Blanco, Madrid-34, Spain.

synthesis through a secondary catalytic event arising from its external effect on a membrane surface component, we interpret this result as meaning that the specific effect of GppCH₂p on picornavirus-infected cells is brought about because this compound is able to pass through the membrane and reach the protein-synthesising machinery only in infected cells. It is thus

suggested that the membrane of all the cells which are actively synthesising picornavirus proteins allows the passage of GppCH₂p and possibly other low molecular weight compounds.

Other nucleotide analogues like AppCH₂p and GppNHp had no significant effect either on infected or uninfected cells

Fig. 1 Specificity of the inhibition of protein synthesis by GppCH₂p. 3T6 cells were grown to confluency in 30-mm Petri dishes containing 2 ml of E4 medium plus 10% foetal calf serum (Difco). The medium was removed and 3 PFU per cell of EMC virus or 0.200 ml E4 added. After 1 h at 37 °C, 1 ml E4 containing 5% calf serum was added back and the dishes incubated for a further 4 h. At that time the medium was replaced by 0.5 ml of E4 medium minus methionine containing 1% calf serum and the indicated amount of GppCH₂p. The cells were pulse labelled with 2 µl [³⁵S]methionine (sp. act. 500 Ci mmol⁻¹; 3 Ci ml⁻¹) at 4½ to 5½ h post-infection (p.i.). The cell monolayer was then washed with phosphate buffer, dissolved in 0.300 ml of appropriate buffer and sonicated. The proteins synthesised were analysed on 15% polyacrylamide gels. 10 µl of sample were applied and the gels run overnight at a constant voltage of 35 V. After staining and destaining the gels were fluorographed and dried. The figure shows the densitometric scans of the autoradiography obtained. The three lower scans on the right-hand side correspond to uninfected 3T6 cells without (control) or with 2 and 1 mM GppCH₂p as indicated. The left-hand side scans correspond to the EMC infected cells (+EMC) minus or plus the indicated concentrations of GppCH₂p. The scans from EMC-infected cells were taken from a more exposed autoradiograph. Or, origin of sample application. The peaks in this and other figures are identifiable from Fig. 4 using the actin (Ac) peak as reference.

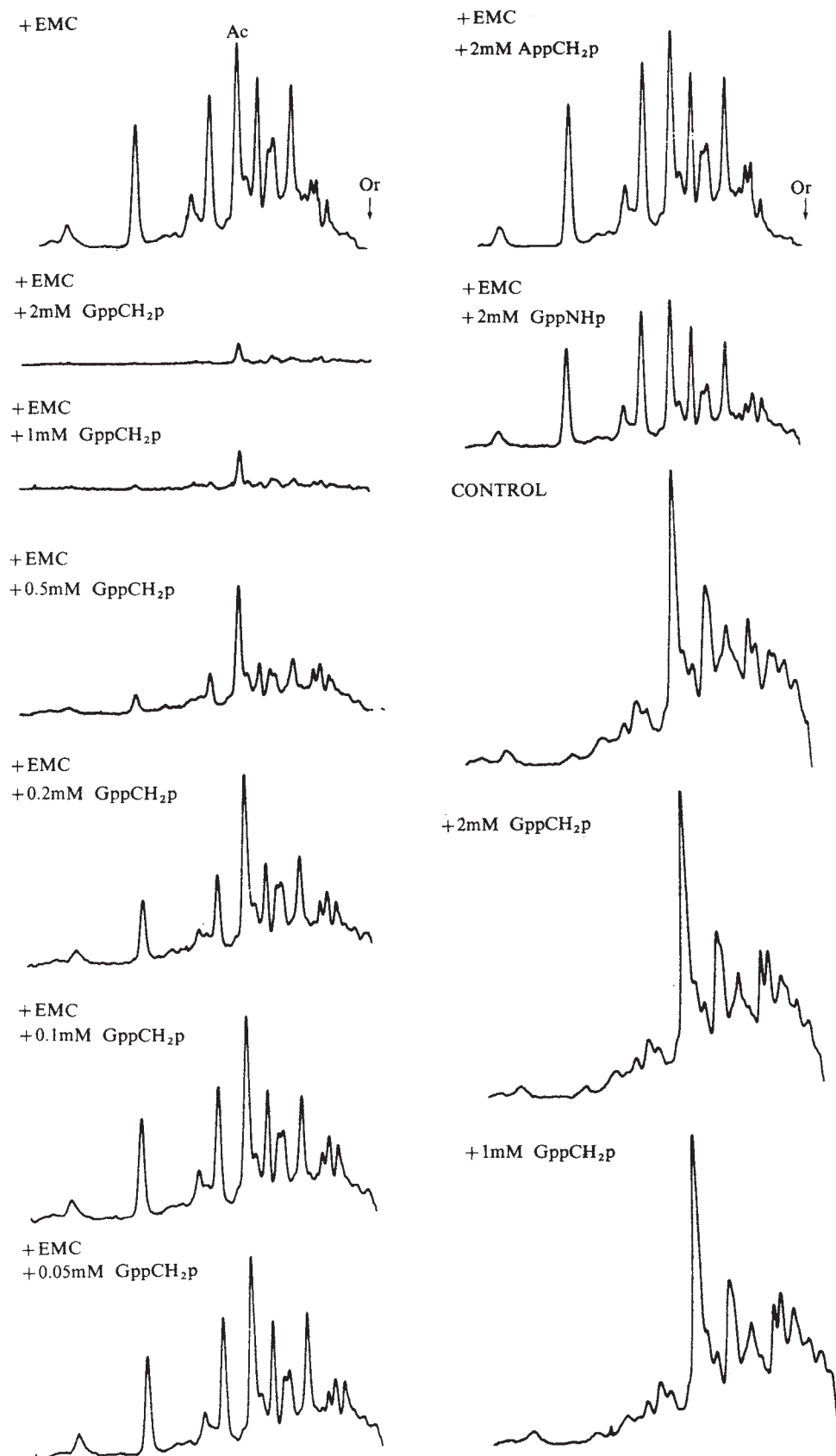


Table 1 Effect of several nucleotide analogues on *in vivo* and *in vitro* protein synthesis and reversal of GppCH₂p inhibition by Gppp

<i>In vivo</i>			<i>In vitro</i>	
Additions	Concentration	[³⁵ S] Methionine incorporated (c.p.m.)	Concentration	[³⁵ S] Methionine incorporated (c.p.m.)
Experiment 1				
None	—	96.229 (100)	—	288.708 (100)
GppCH ₂ p	1mM	27.058 (28)	0.6mM	677 (0.2)
Gppp	1mM	93.068 (97)	1mM	272.512 (94)
Appp	1mM	93.247 (97)	—	—
GppCH ₂ p + Gppp	1mM + 1mM	15.749 (16)	0.6mM + 1mM	2.262 (0.7)
GppCH ₂ p + Appp	1mM + 1mM	14.884 (14)	—	—
Experiment 2				
Control	—	69.112 (100)	—	328.737 (100)
GppCH ₂ p	2mM	15.917 (23)	0.1mM	15.111 (4)
GppNHp	2mM	55.453 (80)	0.1mM	171.000 (46)
AppCH ₂ p	2mM	70.511 (103)	0.5mM	217.260 (58)

The effect of nucleotide analogues *in vivo* was carried out on EMC-infected cells as indicated under Fig. 2a. The nucleotide analogues were added at 4 h p.i. and cells were pulse labelled with ³⁵S methionine at 4½–5½ h p.i. For the *in vitro* experiment the translation of EMC-RNA was carried out in an L cell S-30 system following conditions described elsewhere⁸. The background minus EMC-RNA representing the endogenous incorporation was 65,568 c.p.m. and was not subtracted from the indicated values.

(Figs 1 and 2). This is in agreement with the finding that these two analogues are much less effective inhibitors of protein synthesis *in vitro* (Table 1, ref. 22). In addition, the β-γ amino derivative is also less stable than the β-γ methylene derivative.

Having found a specific inhibitor of protein synthesis in virus-infected cells, we examined the kinetics of this effect. EMC-infected cells were labelled with ³⁵S-methionine for half-hour periods at different times after infection in the presence or absence of GppCH₂p. The results indicate that the nucleotide analogue started to inhibit protein synthesis within 3½–4 h after infection and that this inhibition was concomitant with the appearance of viral proteins (Figs 2 and 3). From this result we suggest that the bulk of viral proteins are synthesised in cells which have a leaky membrane. This finding further supports the membrane-leakage model¹.

Characteristics of the GppCH₂p inhibition

Having observed that there is a temporal relationship between the onset of viral protein synthesis and membrane leakiness, we wondered if gene expression was necessary for the appearance of leakiness, and if so, if the modification was produced by a cellular or a viral function. Exposure of cells to cycloheximide for a period soon after infection reduced the sensitivity to GppCH₂p, implying that gene expression at the translational level was necessary for the induction of membrane leakiness. Evidence that a viral function is involved in making the membrane permeable to GppCH₂p is the observation that the presence of a high concentration of actinomycin D (which suppresses the synthesis of new cellular mRNA but not viral RNA) was without effect on the subsequent inhibition by the nucleotide analogue (Fig. 4). Although unlikely, the possibility that membrane leakiness requires the translation of a masked cellular mRNA which is activated by virus infection is not ruled out by the present results.

Maximal inhibition of protein synthesis was attained by exposing the infected cells to the inhibitor half an hour before the labelling period. This probably indicates that the compound has to diffuse through the membrane to the cytoplasm and reach a certain concentration to prevent translation. The inhibition of protein synthesis by GppCH₂p seems to be a direct effect of this compound on the protein-synthesising machinery, rather than an indirect effect due to the presence of any nucleotide in the culture medium, as neither ATP, GTP (Table 1), AppCH₂p nor GppNHp (Fig. 1) had a similar inhibitory effect. Moreover, the inhibition exerted by GppCH₂p is probably a direct effect of the nucleotide analogue, because it is not reversed *in vivo* or *in vitro* by the addition of GTP (Table 1). This is consistent with the mechanism of action of

GppCH₂p *in vitro* in that it is not a mere competitor with GTP for factors and ribosomes, but rather, once stable complexes have been formed between GppCH₂p and components of the protein-synthesising machinery, they seldom interchange with GTP²³. In addition, the GppCH₂p inhibition is reversed by the addition of fresh medium, arguing against the possibility that its addition to the culture medium triggers the covalent and stable modification of a component of the protein-synthesising apparatus.

Generality and molecular basis of the membrane leakiness

Previous evidence and the results shown here indicate that EMC-infected mouse cells have a leaky membrane late in

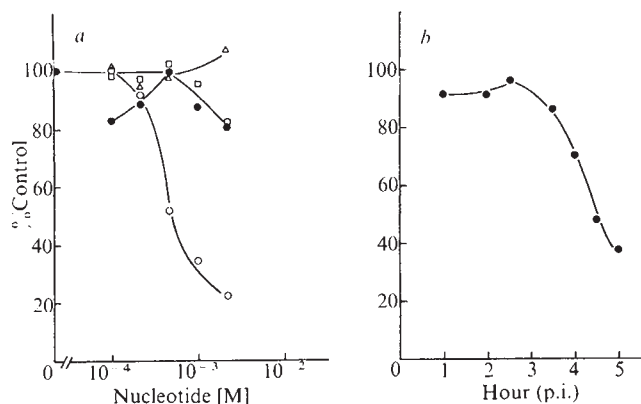


Fig. 2 *a*, Specificity of the inhibition of protein synthesis by GppCH₂p. The cells were infected and processed as indicated in Fig. 1. A 10-μl aliquot was taken after sonication as indicated in Fig. 1 and precipitated with 10% trichloroacetic acid; the sample was then boiled and filtered through glass fibre filters GF/C (Whatman). 100% of the control corresponds to 284,071 c.p.m. of [³⁵S]methionine incorporated into protein for uninfected cells and 114,102 c.p.m. for EMC-infected cells. ●, Uninfected 3T6 cells + GppCH₂p; ○, EMC-infected 3T6 cells + GppCH₂p; □, EMC-infected 3T6 cells + GppNHp; △, EMC-infected 3T6 cells + AppCH₂p. *b*, Time course of the GppCH₂p inhibition. 3T6 cells were grown in 30-mm dishes. All were infected with EMC virus as indicated in Fig. 1. Every 30 min the medium, E4 + 5% calf serum (CS), was replaced by 0.5 ml E4 minus methionine (1% CS), containing 2 μl [³⁵S] methionine and 1 mM GppCH₂p. Controls were infected and labelled in the same way but were not treated with GppCH₂p. 10 μl corresponding to the processed and sonicated samples (see legend to Fig. 1) were filtered as indicated in (*a*). The percentage of the control was calculated by dividing the numbers obtained in the presence of the nucleotide analogue by the corresponding numbers obtained in its absence.

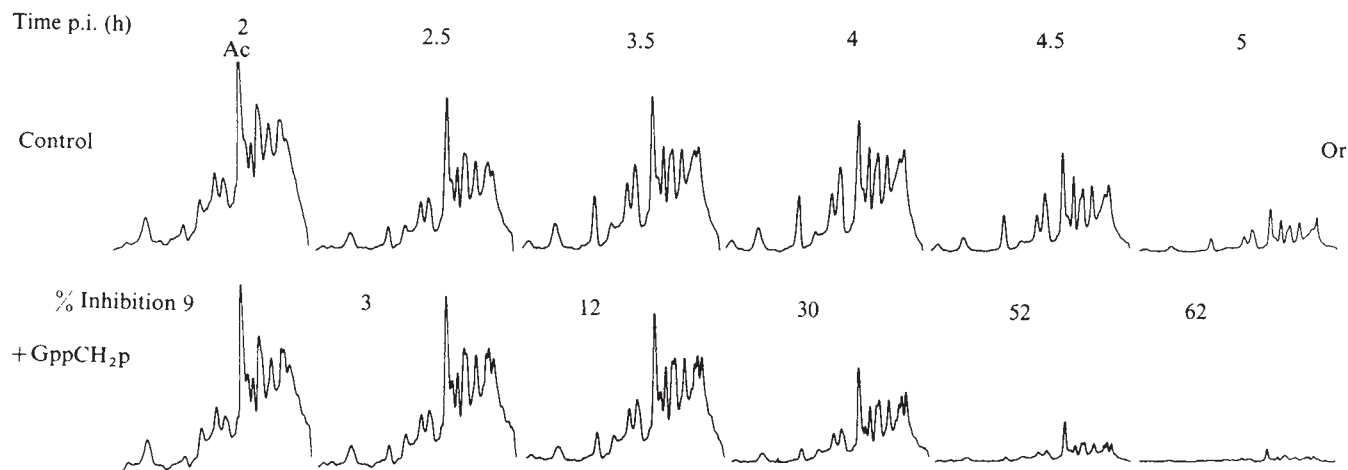


Fig. 3 Time course of the GppCH₂p inhibition. 10- μ l samples from the experiment described in Fig. 2b were analysed on polyacrylamide gels and processed as indicated in Fig. 1. Here we show the densitometric scans obtained from 2 h p.i. onwards. The upper numbers indicate the time p.i. at which the EMC-infected cells were labelled. The first scan then corresponds to the labelling period 2–2½ h p.i. and so on. The lower numbers indicate the % of the inhibition exerted by GppCH₂p.

infection^{5,18}. To determine if this is a general phenomenon of lytic viral infection or if it is a very specific effect observed only with EMC-infected 3T6 cells, we tested the effect of GppCH₂p on other virus-cell systems. Figure 5 shows that Mengo-infected 3T6 and Semliki Forest virus-infected BHK cells behaved like the EMC-infected cells and were specifically inhibited by GppCH₂p. Moreover, we have also observed GppCH₂p inhibition when EMC infects L and BHK cells. These results indicate that impermeable inhibitors can also be used to prevent protein synthesis with viruses other than picornaviruses and with cells other than mouse. Thus, the interference with the cellular membrane may be a widespread phenomenon in infection by lytic viruses.

We have recently tested a number of other protein synthesis inhibitors and found that molecules below 750 molecular weight (MW) enter infected cells, whereas molecules of 10,000–30,000 MW are excluded. These numbers are based on the fact

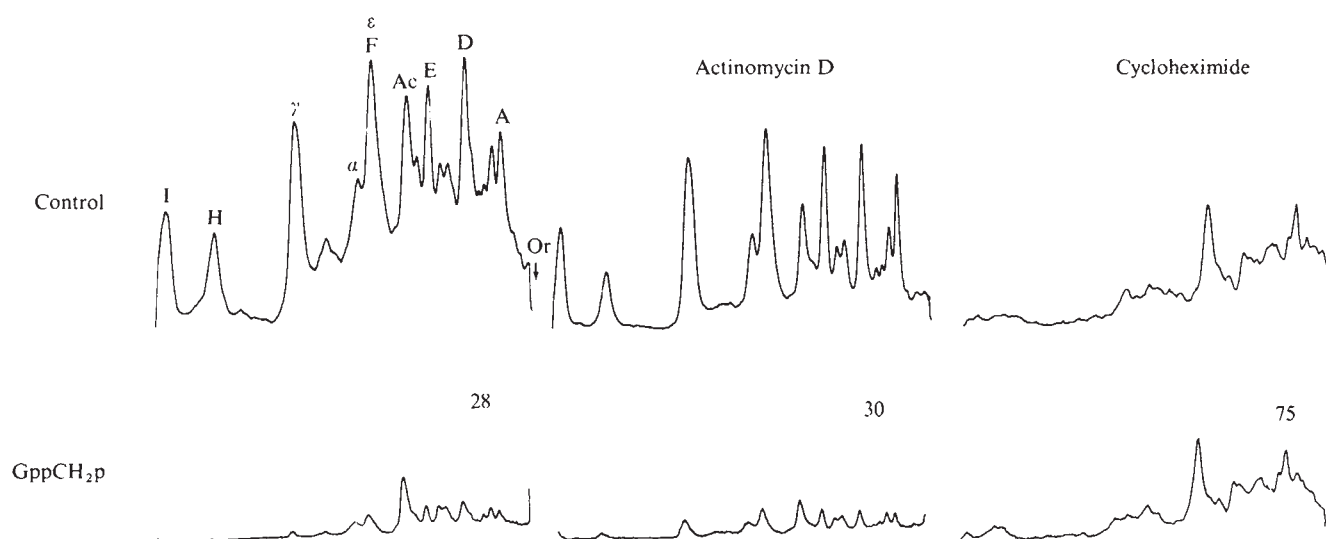
that edeine showed essentially the same effects as GppCH₂p, whereas ricin A chain did not show a preferential inhibition of protein synthesis on virus-infected cells (A. Contreras and L.C., unpublished results).

The lack of specificity in the passage of low molecular weight compounds into infected cells favours the interpretation, recently advanced^{4,24}, that membrane leakiness is caused by holes through which small molecules can pass freely. These holes may be formed by the incorporation of viral proteins in the membrane^{25–27}.

Implications

The previous results showing that the bulk of viral protein synthesis occurs in a cell with a leaky membrane may help to explain some previous findings, the interpretation of which remains obscure. For example, membrane modification, cytopathic effect and cell killing could all be caused by the same

Fig. 4 Gene expression requirement for GppCH₂p inhibition. 3T6 cells were grown in E4 (10% foetal CS). When cells were confluent, the medium was removed and the cell monolayer infected with EMC virus as indicated in Fig. 1. After 1 h at 37 °C, 1 ml E4 (+5% CS) was added and where indicated 5 μ g ml⁻¹ actinomycin D or 10⁻⁴ M cycloheximide. After 4 h the medium was replaced by E4 medium minus methionine (+1% CS), plus or minus 1 mM GppCH₂p, and the cells were labelled as indicated in Fig. 1. All the subsequent operations were carried out as indicated in Fig. 1. The numbers in the lower scans indicate the % of the control. The peaks of the control scan have been labelled to distinguish the different proteins according to the nomenclature of ref. 1.



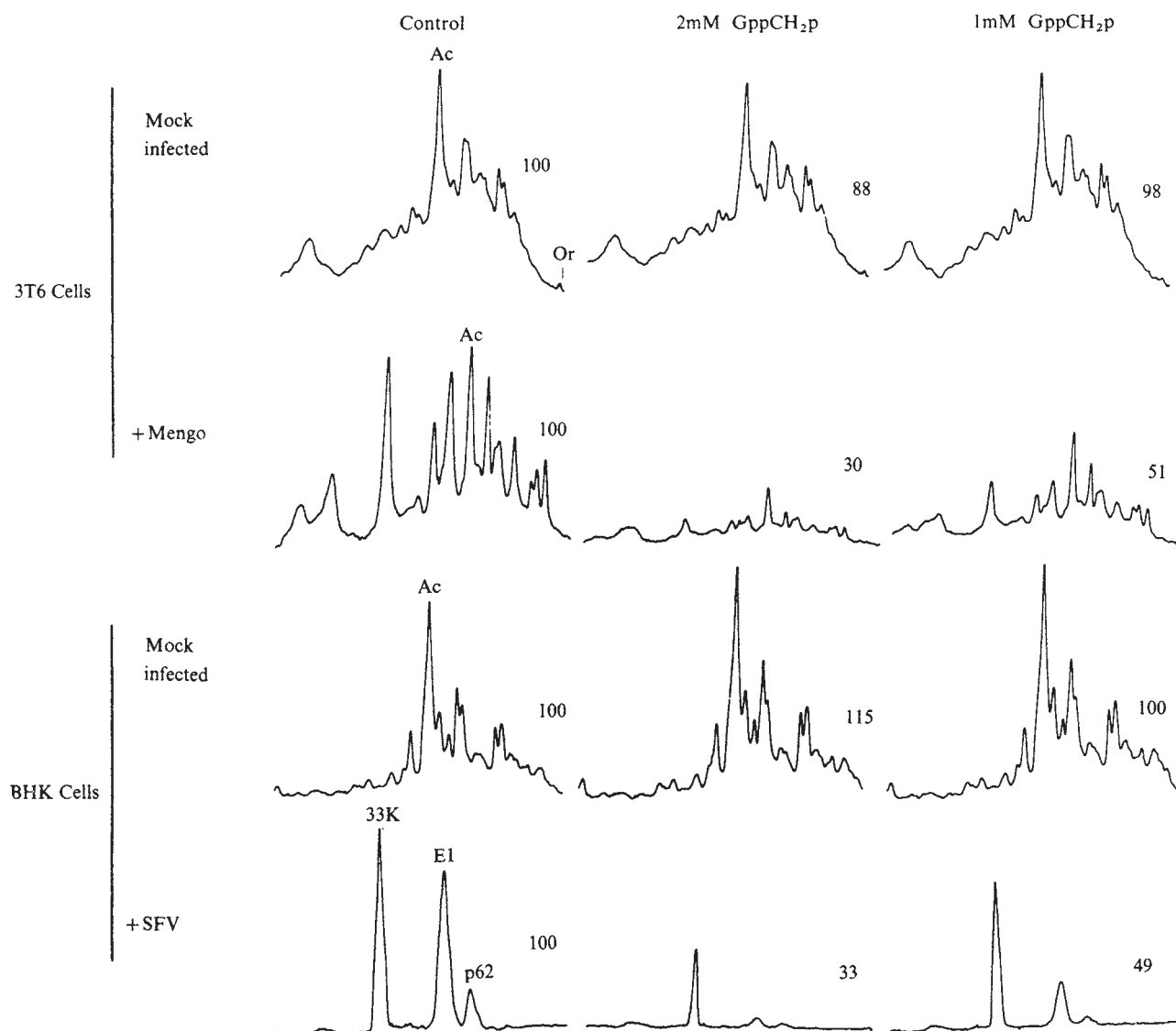


Fig. 5 Specific inhibition by GppCH₂p on Mento and SFV-infected cells. Confluent cultures of 3T6 or BHK C13 cells were mock-infected or infected with Mento virus (5 PFU per cell) or Semliki Forest virus (50 PFU per cell). After addition of fresh medium the cells were kept at 37 °C for 4 h, then the medium was replaced by 0.5 ml E4-methionine (+ 1% CS) containing the indicated amounts of GppCH₂p. Protein synthesis was measured by pulse-labelling the cells with 2 µl ³⁵S-methionine at 5–7 h p.i. The monolayer was processed and the proteins synthesised were analysed as indicated in Fig. 1. The numbers indicate the % of the control. The peaks of the SFV scan are identified according to the nomenclature of ref. 31.

viral protein(s) that is involved in modifying the membrane. In many cases the protein(s) might be a virion coat protein, and if so this may explain reports that a high input of virions causes an alteration of the cellular membrane^{9,10} that in turn influences a spectrum of cellular functions.

These results may also be of use in the design of antiviral agents. To date, one of the major restrictions of antiviral agents is their toxic effect on normal cells. For example, some nucleosides modified in their base moiety are potential carcinogens²⁸. The nucleotide analogue used in this work is in many ways ideal, because to become permeable for normal cells it has to be degraded to the nucleoside, which is unmodified and hence nontoxic. Preliminary experiments indicate that high doses of GppCH₂p are not toxic for mice, and we have observed that GppCH₂p inhibited the production of new infectious virus in tissue culture systems.

It has become an axiom in the design of antiviral agents that they should not interfere with cellular functions because this would be detrimental to the host²⁹. It is generally thought that this requirement makes the finding of a wide spectrum antiviral agent rather unlikely³⁰. However, the approach we propose is the use of inhibitors which interfere with any viral function fol-

lowing the induction of membrane leakiness, irrespective of whether it involves cellular enzymes or not. The important characteristic that antiviral compounds must satisfy is to be impermeable for normal uninfected cells. We believe that inhibitors of protein synthesis will probably have a wider antiviral activity for two reasons: first, all viruses use the same protein-synthesising apparatus³¹ and second, the leakiness of the membrane occurs at about the same time as the onset of synthesis of late viral proteins. Another group of impermeable inhibitors of potential value may be those that interfere with viral assembly itself, because this process would normally occur in a cell with a highly damaged membrane.

I thank Dr A. E. Smith for his help with the manuscript, Drs L. V. Crawford, E. Paucha, E. Rozengurt, P. Rudland and S. Siddell for their critical reading of the manuscript, and the Fundation Juan March for financial support.

Received 24 October 1977; accepted 14 February 1978.

1. Fraenkel-Conrat, H. & Wagner, R. R. (eds) *Comprehensive Virology*, Vols 2–7 (Plenum, New York, 1974–77).
2. Roizman, B. & Spear, P. G. *Curr. Top. Dev. Biol.* **4**, 79–108 (1969).
3. Bablanian, R. *Progr. med. Virol.* **19**, 40–83 (1975).
4. Carrasco, L. *FEBS Lett.* **76**, 11–15 (1977).
5. Carrasco, L. & Smith, A. E. *Nature* **264**, 807–809 (1976).

6. Katzman, J. & Wilson, D. E. *J. gen. Virol.* **24**, 101–113 (1974).
7. Genty, N. *J. Virol.* **15**, 8–15 (1975).
8. Stohman, S. A., Wissman, Ch. L., Jr, Eylar, O. R. & Silverman, D. J. *J. Virol.* **16**, 1017–1026 (1975).
9. Levanon, A., Kohn, A. & Inbar, M. *J. Virol.* **22**, 353–360 (1977).
10. Spiegelstein, M., Fuchs, P., Haimsohn, M., Gitelman, J. & Kohn, A. *J. cell. Biol.* **70**, 414 (1976).
11. Micklem, K. J. & Pasternak, C. A. *Biochem. J.* **162**, 405–410 (1977).
12. Lyles, D. S. & Landsberger, F. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1918–1922 (1977).
13. Norkin, L. C. *J. Virol.* **21**, 872–879 (1977).
14. Herrlich, P., Rahmsdorf, H. J. & Schweiger, M. *Adv. Biosci.* **12**, 523–537 (1974).
15. Condit, R. C. *J. molec. Biol.* **98**, 45–56 (1975).
16. Britton, J. R. & Haselkorn, R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2222–2226 (1975).
17. Ponta, H., Gratzel, M., Pfenning-Yeh, M., Hirsch-Kaufmann, M. & Schweiger, M. *FEBS Lett.* **73**, 207–209 (1977).
18. Farnham, A. E. & Epstein, W. *Virology* **21**, 436–447 (1963).
19. Egberts, E., Hackett, P. B. & Traub, P. J. *J. Virol.* **22**, 591–597 (1977).
20. Ponta, H. *et al. Molec. gen. Genet.* **149**, 145–150 (1976).
21. Stein, W. D. *The Movement of Molecules across Cell Membranes* (Academic, New York, 1967).
22. Eckstein, F., Bruns, W. & Parmeggiani, A. *Biochemistry*, **14**, 5225–5232 (1975).
23. Fernandez-Puentes, C., Carrasco, L. & Vazquez, D. *Biochemistry* **15**, 4364–4369 (1976).
24. Durham, A. C. H. *Nature* **267**, 376 (1977).
25. Hatt, T. & Matthews, R. E. F. *Virology* **73**, 1–16 (1976).
26. Bächli, T., Aguet, M. & Howe, C. J. *J. Virol.* **11**, 1004–1012 (1973).
27. Rifkin, D. B. & Quigley, J. P. *A. Rev. Microbiol.* **28**, 325–351 (1974).
28. Shugar, D. *FEBS Lett.* **40**, S48–S62 (1974).
29. Bauer, D. J. in *Chemotherapy of Virus Diseases* (ed. Bauer, D. J.) **I**, 1–33 (Pergamon, Oxford, 1972).
30. Diana, G. D. & Panic, F. *Angew. Chem., Int. Ed. Engl.* **15**, 410–416 (1976).
31. Smith, A. E. & Carrasco, L. *MTP int. Rev. Sci. Biochem. Ser. II* (in the press).

letters to nature

Faint X-ray sources detected near COS B γ -ray positions

THE Uhuru data has been superimposed with a point summation technique (PST) on the positions of the 11 unidentified COS B γ -ray sources¹. For the other COS B sources, SAS 2 has identified CB185–5 with NP0532 (Crab pulsar) and CG263–2 with the Vela pulsar^{2–4}. We describe here first, the reanalysis by the PST of three weak 4U sources which lie within COS B error circles; second, a survey of the remaining COS B positions which has uncovered a new X-ray source within the CG075+0 error circle; and third, the results of an investigation of the region surrounding CG195+4 and CG189+1.

In the PST, all the Uhuru (1971–72) binned count-rate data scanning a given point of interest on the sky are superimposed to achieve the maximum signal-to-noise ratio. Superimposition is done regardless of collimator orientation; the background is determined on one or both sides of the candidate source position

and the time averaged source intensity and significance are computed on the basis of the expected triangular point source response function. The detection threshold of the 4U catalogue⁵ is ~ 2 Uhuru counts per s (c.p.s.). This catalogue was constructed from multiple candidate sightings each of which is based on ~ 1 d of exposure. In some regions of the Uhuru sky the exposure is sufficient to yield PST intensity uncertainties (1σ) as low as 0.18 c.p.s., which permits a 3σ detection limit of 0.6 c.p.s.

Because of the galactic plane location of the COS B positions, X-ray source confusion severely limits this technique when applied to the 5° collimator data. Three 4U sources have been noted⁵ as lying inside COS B error circles: 4U0241+61/CG135+1, 4U1416–62/CG312–1, and 4U2019+39/CG078+1. Their PST intensities are shown in Table 1 with results for the remaining COS B positions. A new source was detected within the CG075+0 error circle. Its response profile (count rate against scan direction) is shown in Fig. 1a. The profile results from folding the final PST data peak about its

Table 1 Uhuru results for COS B γ -ray sources

COS B source	Intensity* $\pm 1\sigma$	Intensity/ 1σ	PST position†		Comments
			RA	Dec	
064+0	(0) ± 0.33 1.18 ± 0.56	0 2.10	COS B position 301.75 26.50		4 ° from COS B position
075+0	1.60 ± 0.29	5.48	304.28 36.85		
078+1	3.32 ± 0.53	6.25	4U2019+39		Found to be highly variable with PST
121+3	(0) ± 0.59	0	COS B position		
135+1	1.23 ± 0.27 1.38 ± 0.29	4.54 4.76	4U0241+61 40.46 62.08		SAS 3 position
176-7	(0) ± 0.38	0	COS B position		
185-5					Crab pulsar
189+1	3.26 ± 1.18	2.76	92.0	22.2	PST contour position
195+4	1.88 ± 0.70 3.38 ± 0.73	2.68 4.62	98.6 98.5	16.9 17.2	PST contour position 5 ° collimator
263-2					Vela pulsar
312-1	(0) ± 0.36 (0) ± 0.28 2.46 ± 0.33 (0) ± 0.37 (0) ± 0.28 0.31 ± 0.29	0 0 7.44 0 0 1.04	COS B position MX1406-61 4U1416-62 MX1418-61 1L-LEQ87 SNR 208.54 -62.22		PSR1354-62
327-0	0.41 ± 0.53 0.41 ± 0.42 0.40 ± 0.37	0.77 0.96 1.06	COS B position 2S1553-542 PSR1601-52		
333+0	(0) ± 0.56	0	COS B position		

* (0), Intensity obtained was slightly negative but consistent with zero.

† Position at which PST was run. CG195+4 and CG189+1 positions shown are PST contour maxima.

centre to improve statistics. The solid line shows the right-triangle response expected for the intensity obtained. The source has the appearance of a point source and is distinct from nearby Cyg X-3. A light curve shown in Fig. 1b was constructed by summing 5-d intervals with the PST. This curve shows a flare reaching 3.7 Uhuru c.p.s. at the 3.9σ level during 4–10 May 1972, and more activity in early 1973. The change in intensity in May 1972 is about 2.7σ . This sets a light-travel-time upper limit to the size of order 20 light d or $\sim 10^{-3}$ pc.

As CG195+4 has a γ -ray intensity comparable to CG185-5 (Crab Nebula) and has been seen both by SAS 2 (ref. 3) and COS B, an extensive scan of the area surrounding the COS B position was undertaken.

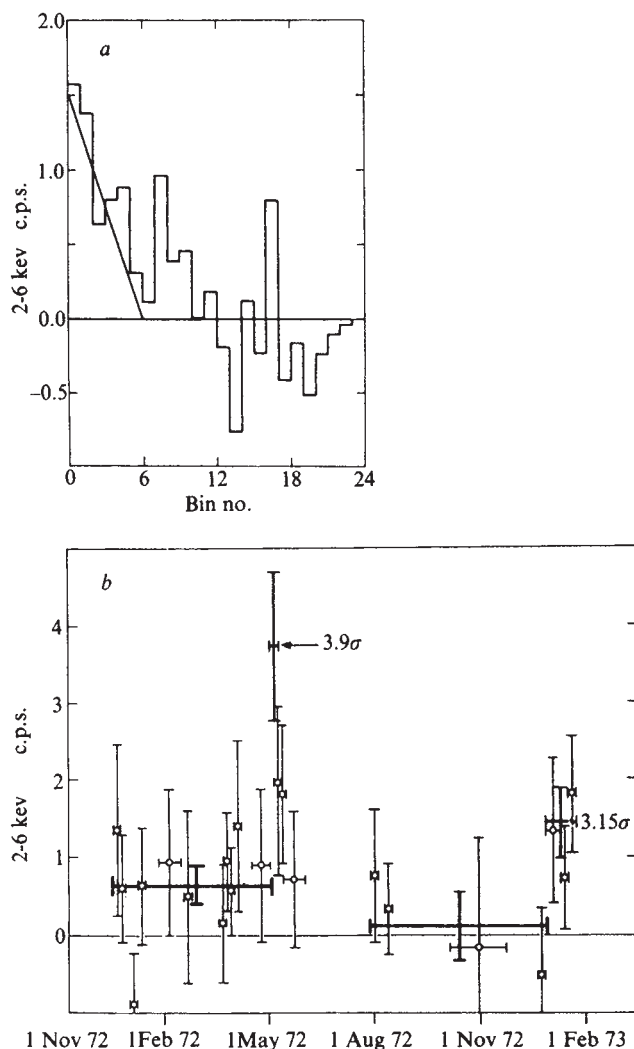


Fig. 1 *a*, Response profile for new X-ray source within CG075+0 error circle. *b*, X-ray light curve for CG075+0 summed on 5-d intervals.

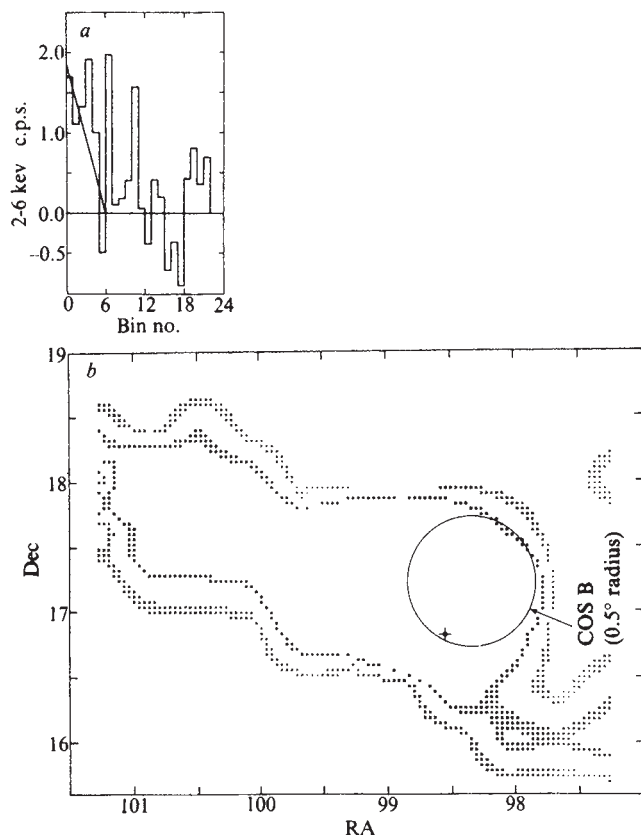
The intensity contour of this region for the narrow collimator data is shown in Fig. 2. The cross shows the position of observed intensity maximum at $\alpha = 98.6^\circ$, $\delta = 16.9^\circ$. The inner contour is that on which the intensity I is equal to the central maximum intensity reduced by twice the standard deviation of this maximum intensity ($I = I_{\max} - 2\sigma_{I_{\max}}$). The outer contour represents the outer limit of positive flux detection. At all points immediately beyond this contour the intensity becomes slightly negative because at these distant locations the source contributes to the background estimate. No attempt has been

made to obtain a rigorous 90% probability contour. These contours, however, indicate a weak Uhuru source localised in the region of CG195+4. As a final check, a PST profile was obtained near the position of the intensity maximum and is also shown in Fig. 2. Again, an approximately triangular response pattern is observed and the source strength is found to be 1.9 c.p.s. at a 2.7σ level of significance. CG195+4 was also seen in the 5° collimator with a roughly consistent intensity and at a confidence level of 4.62σ . Both the Uhuru and COS B error circles are too large to allow any positive source identification with each other.

The available data surrounding CG189+1 were also investigated with the same technique and yielded contour and response profiles shown in Fig. 3. Again, there is evidence for a source at $\alpha = 91.9^\circ$, $\delta = 22.2^\circ$ with intensity of 3.2 c.p.s. at the 2.7σ level of confidence. Finally the contour profiles for CG327-0 and CG064+0 were made but no suggestion of a localised source of X-ray emission was seen within 1° of the COS B position (see Table 1).

The positive results presented here do not seem to be an artefact of the system itself as test runs on ~ 100 arbitrary sky locations have yielded negative results. Furthermore, on the basis of the 4U catalogue only one (or two with the present deeper survey) accidental locations of an X-ray source within a COS B error circle is expected (K. Apparao, personal communication). We find either a 4U source or a probable new X-ray source near 6 of the 11 unidentified COS B sources. We conclude therefore that there is apparently an excess of X-ray sources consistent with COS B γ -ray source positions above that expected by chance. Furthermore, assuming an $E^{-\alpha}$ energy spectrum connecting the X-ray and γ -ray energies, we find that all six detected sources can be fit by an $\alpha = 0.8 \pm 0.1$. More precise positions and spectra for both the X-ray and γ -ray sources are required to establish further connection. Once this

Fig. 2 Response profile (*a*) and intensity contours (*b*) of X-ray emission for CG195+4.



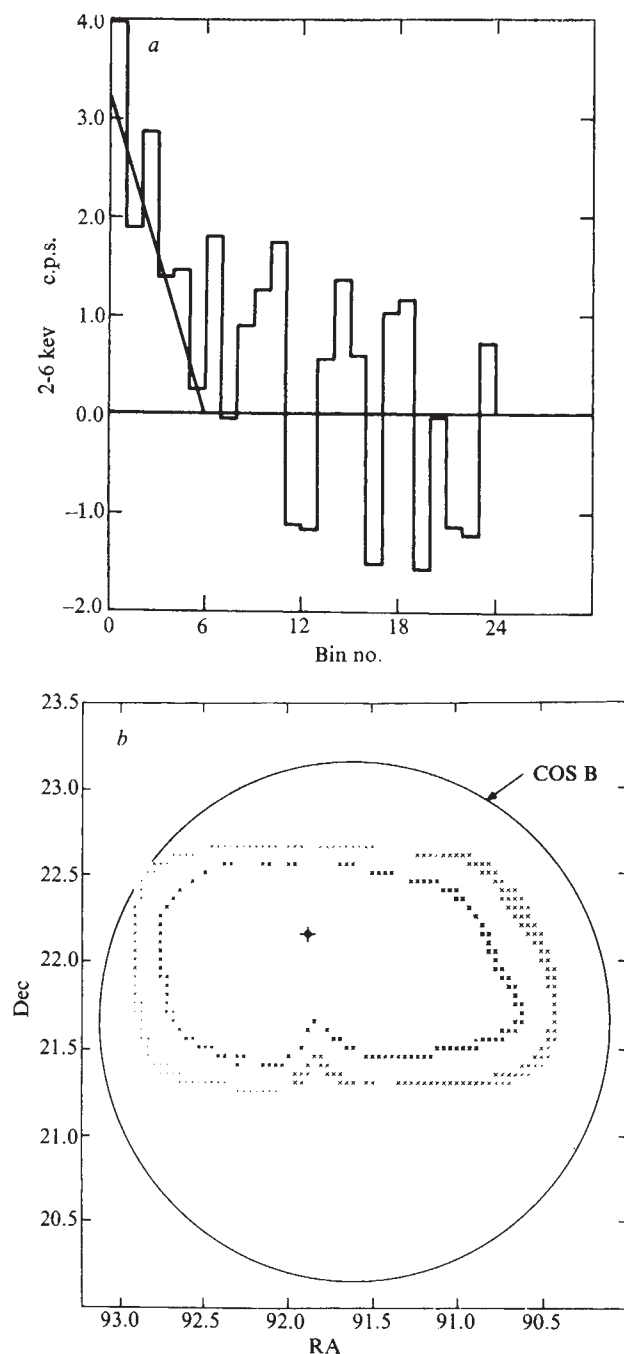


Fig. 3 Response profile (a) and intensity contours (b) of X-ray emission for CG189+1.

has been established, the ability of X-ray astronomy to locate sources to less than $1'$ would allow the optical and/or radio identification of the γ -ray source.

P. F. JULIEN
H. F. HELMKEN

Harvard-Smithsonian Center for Astrophysics,
60 Garden Street,
Cambridge, Massachusetts 02138

Received 13 January; accepted 10 February 1978.

1. Hermesen, W. *et al. Nature* **269**, 494-495 (1977).
2. Thomson, D. J., Fichtel, C. E., Kniffen, D. A. & Ogelman, H. B. *Astrophys. J.* **200**, L79-L82 (1975).
3. Thomson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252-262 (1977).
4. Lamb, R. C., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Thomson, D. J. *Astrophys. J.* **212**, L63-L66 (1977).
5. Maraschi, L. *et al. Nature* **272**, 679-681 (1978).
6. Forman, W. *et al. Astrophys. J. Suppl.* (submitted).

Positions of galactic X-ray sources: $-20^\circ < l'' < +6^\circ$

THE precise (20 – $60''$) positions of nine X-ray sources, in the vicinity of the galactic centre are reported here. The data were obtained as part of the comprehensive survey of the galactic plane performed with the rotating modulation collimator (RMC) detectors on the SAS-3 X-ray observatory^{1–6}. One of the nine is the binary X-ray source 4U1700–37 which has a well established optical counterpart that lies $7''$ from the position reported here. This supports the validity of our calibration procedures². The other eight sources GX349+2, 4U1702–42, 4U1705–44, MX1716–31, A1742–294, 4U1755–33, GX5–1, and 2S1803–245 lack established counterparts at other wavelengths. Our position for GX5–1 adds confidence to the previously proposed radio counterpart⁷ and our position for 4U1755–33 excludes its proposed optical counterpart⁸. Three of the unidentified sources have distinguishing X-ray properties: MX1716–31 is a flaring source^{9,10} and possibly an X-ray burst source¹¹; A1742–294 is a transient¹² and possibly also the X-ray burst source MXB1742–29 (ref. 13); and 2S1803–245 (MX1803–24) is a very bright transient discovered¹⁴ in the course of this work.

Two or more independent sets of position measurements were made with the $2.3'$ and $4.5'$ FWHM rotating modulation collimators (RMC) for all the sources except MX1716–31 and 2S1803–245. Each independent observation had a typical effective exposure of 2×10^4 s. The sources MX1716–31 and 2S1803–245 were detected in two recent observations with deep exposures of $\sim 10^5$ s.

During the observation of MX1716–31 the RMC system reached an unusually low temperature of -20°C which is significantly below the range encountered during our earlier observations². Three calibration observations of Sco X-1 were undertaken to extend the calibrations to lower and higher temperatures and thereby to obtain temperature-dependent calibration parameters. In this work, only the position of MX1716–31 depends solely on these parameters. We have also located an error in the counting statistics term, σ_1 , in our expression² for the radius of the error circle to be associated with a measured celestial position. The correct term is $\sigma_1 = 0.87\delta N_T^{1/2} N_S^{-1}$. The earlier version gave an error radius which was too large by an amount typically $\leq 10''$.

The positions, error radii, and intensities (2 – 11 keV) of all nine sources are given in Table 1. The SAS-3 error circles are compared to other position determinations in Fig. 1. Finding charts for all sources except 4U1700–37 are presented in Fig. 2. Precise stellar positions for these charts are given in Table 2.

We summarise below the status of each source with regard to its distinguishing X-ray properties and any proposed optical and radio counterparts:

4U1700–37 (2S1700–377): The well established optical counterpart of this source exhibits a 3.4-d periodicity^{15,16} that was previously discovered¹⁷ at X-ray wavelengths.

GX349+2 (2S1702–363): This bright X-ray emitter, typical of sources in the galactic bulge, has no distinguishing properties other than its high luminosity.

4U1702–42 (2S1702–429): The source was detected in the present work at a flux density comparable to the Uhuru flux density.

4U1705–44 (2S1705–440): The flux density detected by SAS-3 is a factor of five less than that given in the Uhuru catalogue¹⁸. This confirms the reported high variability of this source¹⁸.

MX1716–31 (2S1715–321): This source was first detected by the OSO-7 satellite as a steady emitter of X rays⁹. Subsequently, it was found to emit intense flares of a few minutes duration with an X-ray flux ~ 60 times

Table 1 Celestial positions

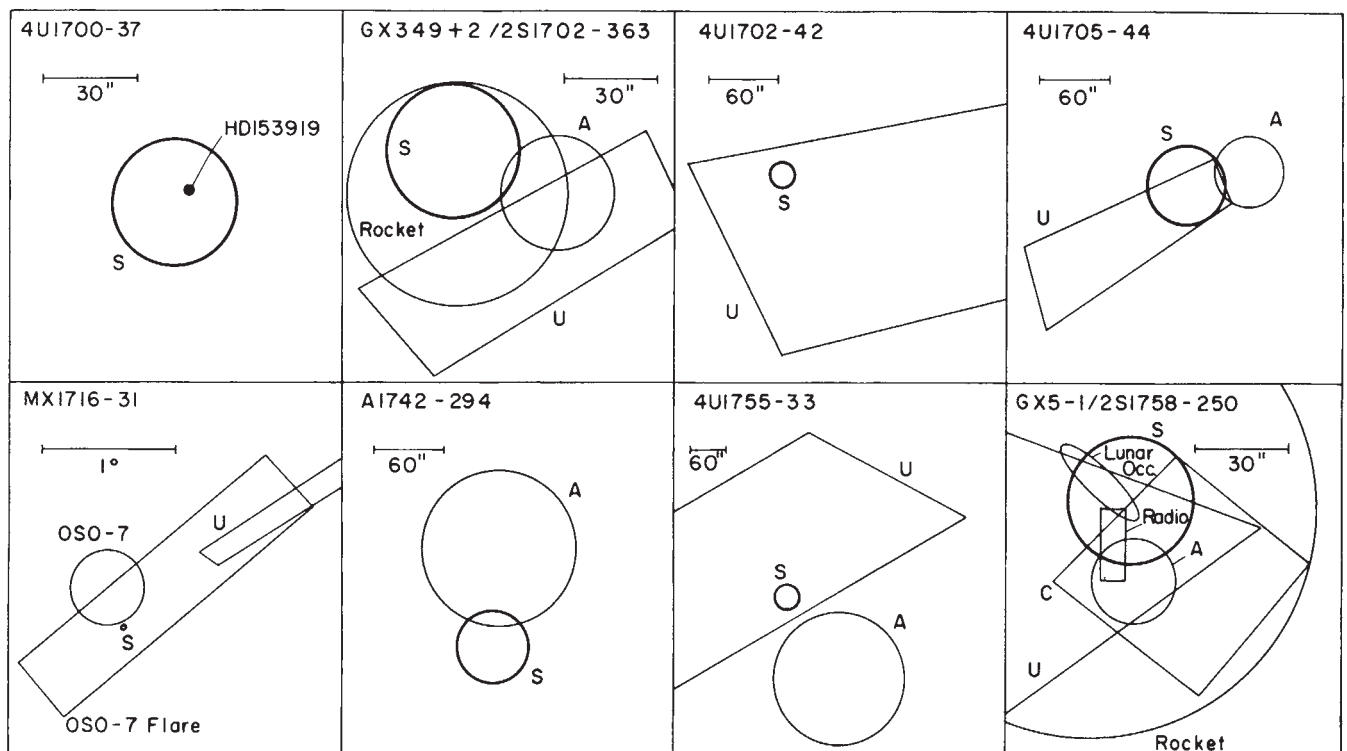
SAS-3 designation	Other designations*	α	Position (1950) δ	l^{11} b^{11}	Error radius (90%)	Flux density† (2–11 keV)	Comments
2S1700–377	4U1700–37	17 h 00 m 33.1 s 255.1379°	–37° 46′ 33″ –37.7758	347.8 2.2	20″	73 μ Jy (June–July 1975)	Optical identification ^{15–17}
2S1702–363	GX349+2 4U1702–36	17 02 23.3 255.5971	–36 21 23 –36.3564	349.1 2.7	20	581 (May–July 1975, May 1977)	
2S1702–429	4U1702–42	17 02 40.4 255.6683	–42 57 58 –42.9661	343.9 –1.3	30	50 (June 1975)	
2S1705–440	4U1705–44	17 05 18.2 256.3258	–44 02 10 –44.0361	343.3 –2.3	20	51 (June 1975)	
2S1715–321	MX1716–31 4U1704–30?	17 15 31.9 258.8829	–32 07 15 –32.1208	354.1 3.1	60	27 (May 1977)	Flaring source ¹⁰ Fast transient ¹¹
2S1742–294	A1742–294 A1743–29 GCX	17 42 53.6 265.7233	–29 29 50 –29.4972	359.6 –0.4	30	58 (Aug. 1975, May–June 1977)	Transient ¹² MXB1742–29 ¹³
2S1755–338	4U1755–33	17 55 22.1 268.8421	–33 48 25 –33.8069	357.2 –4.9	20	63 (June 1975, May 1977)	
2S1758–250	GX5–1 3U1758–25	17 58 3.0 269.5125	–25 04 39 –25.0775	5.1 –1.0	20	862 (Aug. 1975)	Radio candidate ⁷
2S1803–245	MX1803–24	18 03 45.8 270.9408	–24 35 38 –24.5939	6.1 –1.9	30	850‡ (May 1976)	Transient ¹⁴

* Refs 9, 12, 14, 18, 22, 23, 28.

† Averaged over all observations. 1.0 μ Jy corresponds to 2.2×10^{-11} erg cm⁻² s⁻¹ (2–11 keV). $I_{\text{crab}} = 1,060$ μ Jy (see refs 1, 2).

‡ Maximum observed flux density May 16–19 1976.

Fig. 1 The SAS-3 error regions for source locations are compared with other measurements. (Separate scale factors are indicated for each source). Positions are from the 4U catalogue (U, ref. 18), the Ariel V rotating modulation collimator (A, ref. 22), the Copernicus satellite (C, ref. 25), the OSO-7 satellite (refs 9, 10), the MIT rocket-borne RMC (refs 23, 28), and the SAS-3 RMC system (S (dark circles), present work). The optical counterpart of 4U1700–37, HD153919, is located well within the SAS-3 error region. The OSO-7 location for the flare event from MX1716–31 is consistent with the position of the steady source. The lunar occultation error region for GX5-1 and the position of its candidate radio source are presented. The source 2S1803–245 is not displayed as no other measurements are available. North is up and east is to the left.



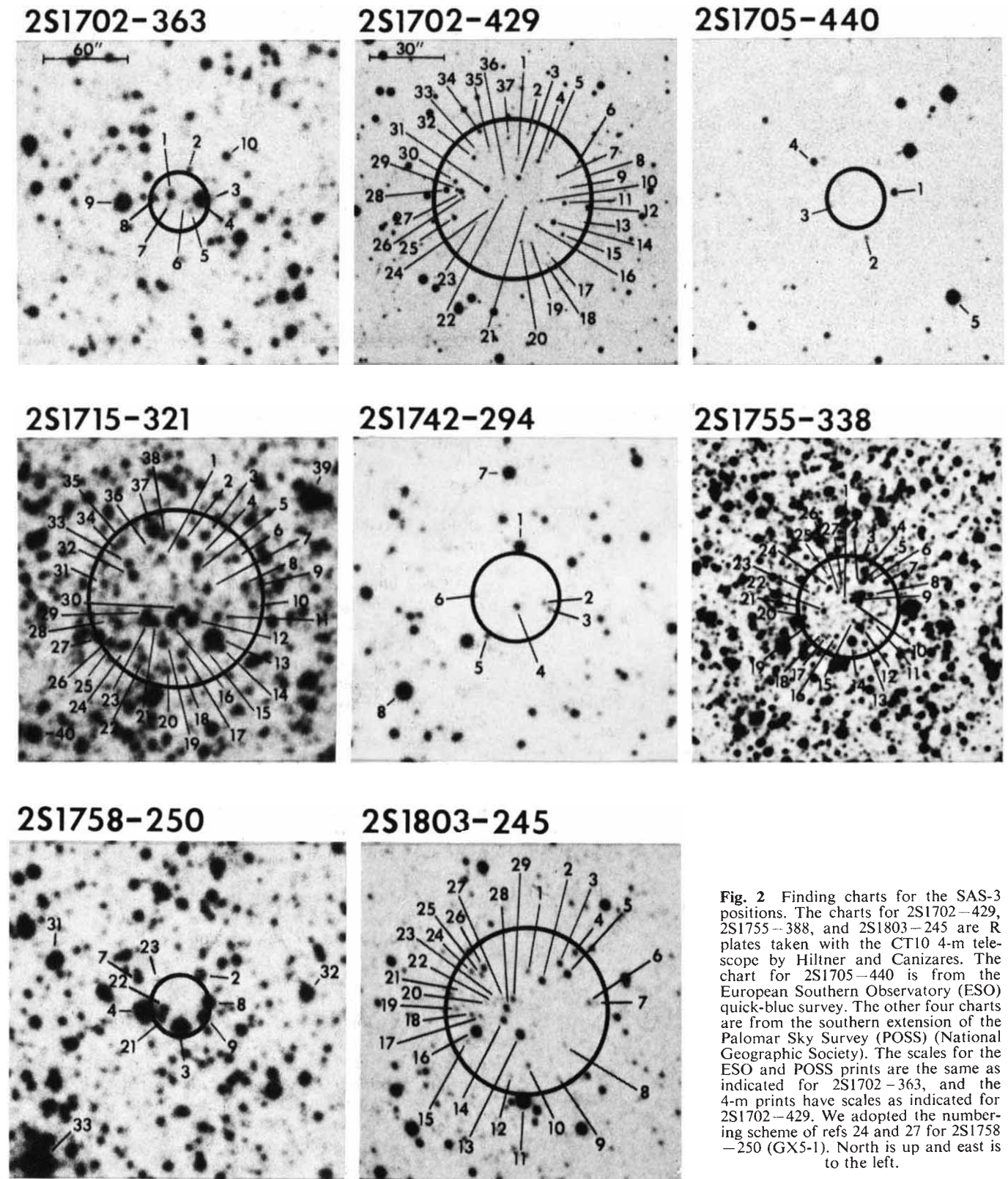


Fig. 2 Finding charts for the SAS-3 positions. The charts for 2S1702-429, 2S1755-388, and 2S1803-245 are R plates taken with the CT10 4-m telescope by Hiltner and Canizares. The chart for 2S1705-440 is from the European Southern Observatory (ESO) quick-blue survey. The other four charts are from the southern extension of the Palomar Sky Survey (POSS) (National Geographic Society). The scales for the ESO and POSS prints are the same as indicated for 2S1702-363, and the 4-m prints have scales as indicated for 2S1702-429. We adopted the numbering scheme of refs 24 and 27 for 2S1758-250 (GX5-1). North is up and east is to the left.

greater (3–10 keV) than the steady flux¹⁰. The present work yielded an X-ray intensity comparable to the steady OSO-7 intensity. In an independent SAS-3 observation, an unusually fast transient X-ray event lasting ~100 s was detected with the RMC system¹¹. The intensity was modulated by both of the modulation collimators and the results jointly restrict the possible source locations to <1% of the common 12° × 12° FWHM field of view. The only known steady X-ray source which lies in the allowed regions is

MX1716-31. It is possible, though far from certain, that the ~100 s burst came from MX1716-31.

A1742-294 (2S1742-294): This transient source was discovered by the Ariel V group¹² and is located within the complex X-ray emitting region, GCX, in the immediate vicinity of the galactic centre¹⁹. This source is one of several^{12,13,20,21} which form this complex. It is the only source in GCX resolved in the present analysis. The same region contains at least three X-ray burst sources of which

one, MXB1742-29 (ref. 13), may be identified with A1742-294.

4U1755-33 (2S1755-338): A suggested optical candidate SAO209568 (ref. 8) lies outside the Ariel V (ref. 22) and SAS-3 error regions. No other optical or radio candidates have been reported.

GX5-1 (2S1758-250): This source is the brightest member of the galactic bulge X-ray sources. A radio source has been suggested as a counterpart⁷. Precise positions, including one derived from lunar occultations, have been reported by several groups²²⁻²⁵. Together, these positions support the identification with this radio source. Searches for infrared²⁶ and optical^{26,27} counterparts have been inconclusive.

2A1803-245: This transient source was discovered with SAS-3 in May 1976 (ref. 14). It was first detected on 5 May with the slit detectors at a flux density of $\sim 100 \mu\text{Jy}$ (2-11 keV). It was then observed with the RMC detectors 12-16 May during which time it increased in brightness from $\sim 320 \mu\text{Jy}$ to $\sim 850 \mu\text{Jy}$ (2-11 keV). It was detected again for the last time on 19 May at $\sim 850 \mu\text{Jy}$ (2-11 keV).

Table 2 Stellar astrometry

Source	Star	Position (1950)*			
		α	δ		
2S1702-363	9	17 h 02 m 26.6 s	$-36^\circ 21' 23''$		
	10	17 02 20.6	$-36 20 48$		
2S1705-440	4	17 05 21.0	$-44 01 44$		
	5	17 05 11.6	$-44 03 19$		
2S1715-321	39	17 15 24.4	$-32 06 03$		
	40	17 15 40.1	$-32 08 50$		
2S1742-294	7	17 42 54.5	$-29 28 22$		
	8	17 42 59.7	$-29 30 56$		
2S1758-250	3	17 58 02.9	$-25 04 54$		
	31	17 58 09.4	$-25 04 09$		
	32	17 57 56.5	$-25 04 31$		
	33	17 58 10.1	$-25 06 25$		
2S1803-245	11	18 03 45.7	$-24 36 12$		

* Precise to $\lesssim 3''$.

We thank the staffs of the Center for Space Research at MIT and the Goddard Space Flight Center for their support, also R. Kelley, J. Kulik, E. Ng, and C. Reid for their assistance. We thank M. Liller, R. McCrosky, and C. Shao of the Harvard College Observatory for facilities for measuring stellar positions. This work was supported in part by the US NASA under contract NAS5-11450. This article is the seventh in a series on positions of galactic X-ray sources obtained with SAS-3.

J. G. JERNIGAN
K. M. V. APPARAO*
H. V. BRADT
R. E. DOXSEY
R. G. DOWER
J. E. MCCLINTOCK

Department of Physics and Center for Space Research,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

Received 29 December 1977; accepted 6 February 1978.

* Permanent address: Tata Institute of Fundamental Research, Bombay, India.

- Bradt, H. V. *et al.* *Nature* **269**, 21-25 (1977).
- Doxsey, R. E., Apparao, K. M. V., Bradt, H. V., Dower, R. G. & Jernigan, J. G. *Nature* **269**, 112-116 (1977).
- Bradt, H. V. *et al.* *Nature* **269**, 496-497 (1977).
- Jernigan, J. G., Apparao, K. M. V., Bradt, H. V., Doxsey, R. E. & McClintock, J. E. *Nature* **270**, 321-323 (1977).
- Doxsey, R. E., Apparao, K. M. V., Bradt, H. V., Dower, R. G. & Jernigan, J. G. *Nature* **270**, 586-588 (1977).
- Apparao, K. M. V., Bradt, H. V., Dower, R. G., Doxsey, R. E., Jernigan, J. G. & Li, F. *Nature* **271**, 225-228 (1978).
- Braes, L. L. E., Miley, G. K. & Schoenmaker, A. A. *Nature* **236**, 392 (1972).
- Jones, C. A., Chetin, T. & Liller, W. *Astrophys. J. Lett.* **190**, L1-L3 (1974).
- Markert, T. H. *et al.* *Astrophys. J.* **218**, 801-814 (1977).
- Markert, T. H., Backman, D. E. & McClintock, J. E. *Astrophys. J. Lett.* **208**, L115-L118 (1976).
- Hoffman, J. A. *et al.* *Astrophys. J. Lett.* (in the press).
- Ariel V Group, *IAU Circ.* No. 2934 (1976).
- Lewin, W. H. G. *et al.* *Mon. Not. R. astr. Soc.* **177**, 83p-92p (1976).
- Jernigan, J. G. *IAU Circ.* No. 2957 (1976).
- Penny, A. J., Olwin, R. P., Penfold, J. E. & Warren, P. R. *Mon. Not. R. astr. Soc.* **163**, 7p-12p (1973).
- Hutchings, J. B., Thackeray, A. D., Webster, B. L. & Andrews, P. J. *Mon. Not. R. astr. Soc.* **163**, 13p-18p (1973).
- Jones, C. *et al.* *Astrophys. J. Lett.* **181**, L43-L48 (1973).
- Forman, W. *et al.* *Astrophys. J. Suppl.* (in the press).
- Kellogg, E., Gursky, H., Murray, S., Tananbaum, H. & Giacconi, R. *Astrophys. J. Lett.* **169**, L99-L103 (1971).
- Eyles, C. J., Skinner, G. K. & Willmore, A. P. *Nature* **257**, 291-293 (1975).
- Criddle, R. G., Fritz, G., Shulman, S., Friedman, H., McKee, J. & Johnson, M. *Astrophys. J. Lett.* (in the press).
- Wilson, A. M., Carpenter, G. F., Eyles, C. J., Skinner, G. K. & Willmore, A. P. *Astrophys. J. Lett.* **215**, L111-L115 (1977).
- Schnopper, H. W. *et al.* *Astrophys. J. Lett.* **161**, L161-L167 (1970).
- Hoffman, J. A., Davison, P. J. N. & Morrison, L. V. *Nature* **244**, 347-349 (1973).
- Willmore, A. P. *et al.* *Mon. Not. R. astr. Soc.* **169**, 7-23 (1974).
- Murphy, P. *et al.* *Mon. Not. R. astr. Soc.* **169**, 25-34 (1974).
- Kunkel, W. *et al.* *Astrophys. J. Lett.* **161**, L169-L172 (1970).
- Rappaport, S., Zaamen, W., Doxsey, R. & Mayer, W. *Astrophys. J. Lett.* **169**, L93-L97 (1971).

New highly variable radio source, possible counterpart of γ -ray source CG135+1

THE discovery of a highly variable radio source which coincides in position with cosmic γ -ray source CG135+1 (ref. 1) is reported here. The radio source, which we designate GT 0236+610, was discovered during a survey of the galactic plane for a class of highly variable radio sources. The survey was aimed at the detection of sources whose radio emission is normally very weak or absent but which from time to time exhibit short periods of intense radio emission. The prototype of this class of object is the X-ray source Cyg X-3 which was found² to exhibit giant radio outbursts of duration approximately 1 week at irregular intervals of a few months. The survey consisted of repeatedly mapping an area of approximately 50 deg^2 along the galactic plane, every day for 19 d. The daily results were subtracted from the average of the 19 d for evidence of variability.

The observations were carried out with the NRAO 91-m transit telescope at a wavelength of 6 cm between 11 and 30 August 1977. The mapping operation was achieved by driving the telescope in declination at the maximum servoed rate of $130' \text{ min}^{-1}$ back and forth across the galactic plane, over the RA range $19 \text{ h } > 3 \text{ h}$ ($l = 40^\circ \rightarrow 140^\circ$). Each declination scan was 4° long centred on the galactic plane and consecutive scans were displaced in right ascension by an amount $\Delta \text{RA} = 2 \text{ min } 20 \text{ s}$ as a result of Earth rotation. The telescope is equipped with a rotatable dual feed system which produces two beams, each with a HPBW = $2.7'$, which are displaced by $3.7'$ on either side of the telescope axis of symmetry. The receiver system consists of two separate 6-cm cooled receivers each with a system temperature of 66 K and a 3-db bandwidth of 580 MHz. The two feed horns are switched rapidly between the inputs of the two receivers in such a fashion that when feed A is connected to receiver 1, feed B is connected to receiver 2 and vice versa. This yields two independent outputs which in addition to providing a $\sqrt{2}$ improvement in sensitivity greatly increases the reliability of the detection of a transient source.

For each scan the two beams were orientated in position angle so that the tracks of the two beams on the celestial sphere were separated by $\frac{1}{2}$ HPBW. An unresolved source midway between the two tracks produces an 'S' shaped instrumental profile with equal positive and negative peaks, whereas a source offset with respect to the centre of the two tracks produces unequal + and - peaks. The ratio of the two peaks is independent of source strength. From this ratio and a knowledge of the beam shapes it is possible to determine the offset or position of the source. By this means we were able to distinguish true source variability from possible variations produced by small changes in the telescope pointing from day to day. The beam shapes were carefully measured

using calibration sources spread out over the declination range covered in the survey, to allow for changes in beam shape with declination.

One of our principle concerns before carrying out this survey was the problem of detecting weak variable sources in the presence of relatively strong confusing sources such as the HII regions, found in the galactic plane. The effect of beam switching with two closely spaced beams reduces this problem considerably by filtering out sources with an angular size significantly larger than $7'$. Great care was also taken in the calibration of the receiver and telescope gain. The receiver gains were calibrated once every 5 min using a stable noise tube. In addition three calibration sources 3C48, DR21 and NGC7027 were observed each day, with scans identical to those used in the survey, to monitor the telescope gain and the stability of the telescope pointing. As a further precaution, the observations were carried out at night to avoid possible variable solar heating effects on telescope gain and pointing.

Table 1 summarises the results from our three calibration sources. The quantities tabulated are those capable of producing significant apparent source variations. Column *A* lists the r.m.s. pointing variations in the direction perpendicular to the sky track defined by the telescope motion. The related source variations are a function of both the magnitude of this r.m.s. and the actual value of the source offset with respect to the track. A maximum (r.m.s.) source variation of approximately 5% could occur if the source was situated on the steepest part of the beam profile. Pointing changes along the track have a typical r.m.s. of $14''$ but because of the method of analysis can produce artificial source variations no greater than 0.5%. Column *B* shows the r.m.s. of the total observed source variations, including the effects of both telescope gain and pointing. From columns *A* and *B* and a knowledge of the beam profiles, the telescope gain variations can be derived and these are listed in column *C*.

Table 1 Observational results on 3C48, DR21 and NGC7027

Source	<i>A</i> Observed r.m.s. variations in source offset (arc s)	<i>B</i> Total r.m.s. observed source variation (%)	<i>C</i> Telescope r.m.s. gain variation (%)
3C48	1.2	2.0	1.7
NGC7027	2.4	2.8	1.8
DR21	3.2	2.3	0.8

Figure 1a and b shows the raw data for GT 0236+610 for both receivers on two days, one when the source was strong and another when it was almost completely absent. The r.m.s. fluctuations on an individual scan correspond to 15 mK antenna temperature.

The agreement between the scans on the two different days is very close with the exception of a clear source near the beginning of the scan on 27 August. An average scan was constructed from days when the source was not apparent and each day's result subtracted from this average. Figure 3c shows the average difference for both receivers for 27 August when the source was strongest. From the ratio and the location of the positive and negative peaks we have derived the following position for the source: RA (1950) = 02 h 36 min 41 s $\pm$ 8 s; dec (1950) = $61^\circ 01' 24'' \pm 30''$.

We have considered the possibility that our variable source might be produced by a strong source located at the edge of the beam. If this were the case then the variations in the telescope pointing would need to be large enough to produce an increase of at least a factor of 10 in the observed source strength. In contrast the r.m.s. source variation expected due to telescope pointing errors for a source on the steepest portion of the beam is only 5%. Furthermore the nearest catalogued

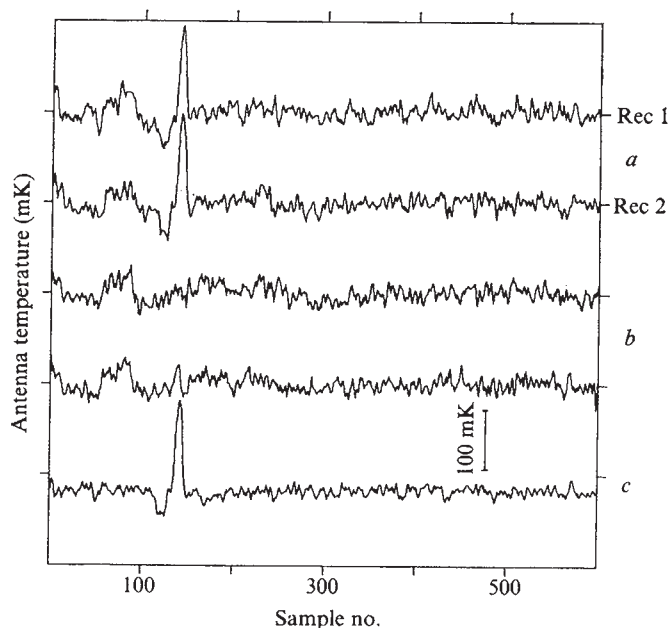


Fig. 1 The output of both receivers when the source is (a) strong, on 27 August, and (b) weak, on 15 August. (c) The signal on 27 August after combining the outputs of both receivers and subtracting an average scan constructed from days when the source is not apparent.

radio source 4C61.06 is $13'$ away at RA (1950) = 02 h 36 min 41.3 s; dec (1950) = $61^\circ 14' 12''$ and is relatively weak with a flux density of 2.2 Jy at 179 MHz. We therefore conclude that GT 0236+610 is a true variable source.

Figure 2 shows the observed flux density of the source for the entire observing period. The most prominent feature is a flare which commenced on or about 25 August and reached a maximum on 27 August, thereafter slowly subsiding. Before the flare the source appeared to be variable within the range < 15 mJy to ~ 75 mJy. The rise time of the flare places an upper limit on the source radius of approximately 1 light day = 2.5×10^{15} cm, however, with our one day sampling interval we cannot exclude the possibility of shorter time variations. From the upper limit to the source size and the observed maximum flux density, we can deduce a lower limit to the source brightness temperature T_b as a function of the assumed distance (D). $T_b(K) \geq 1.4 \times 10^5 D^2$ (D is in kpc).

Radio sources which exhibit variability similar in character to GT 0236+610 are exceedingly rare and all previously known examples are identified with binary stars (such as, Cyg X-3 (ref. 2), Cir X-1 (ref. 3) and Algol (ref. 4)). This fact combined with the galactic latitude of the source $b^{11} = 1^\circ$ argues strongly that the radio source is galactic and probably associated with a binary star. We have examined the sky survey prints for a possible optical candidate. The brightest object in the radio error box is listed in the catalogue (see ref. 5) as LSI +61°303 and described as an OB+ star. *UBV* photometry of this star by Drilling⁶ yields $V = 10.79$ mag, $B - V = 0.8$ mag and $U - B = 0.33$ mag from which we derive a visual extinction $A_v = 3.3$ mag. Herr's⁷ results indicate that 46% of the stars classified as nonemission OB+ in these catalogues are supergiants with absolute visual magnitude ranging from -6 to -7 . This gives a distance to the star of 5–8 kpc which would place it close to the edge of the galaxy for this galactic longitude.

Figure 3 shows the location of the radio source within the error box of the COS B γ -ray source CG135+1 (ref. 1). Also plotted is the location of the only known X-ray source in this region 4U0241+62 for which an accurate SAS-3 position was recently

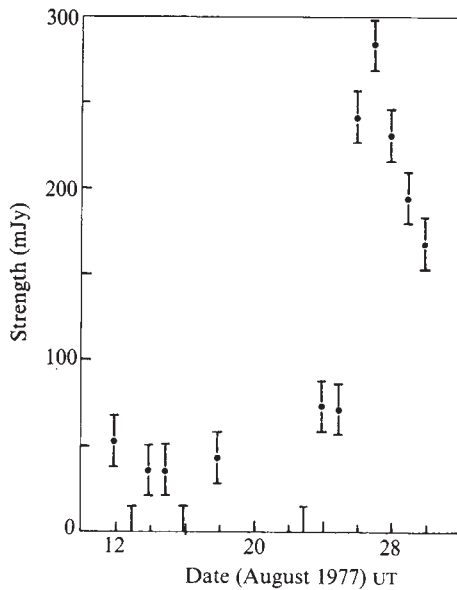
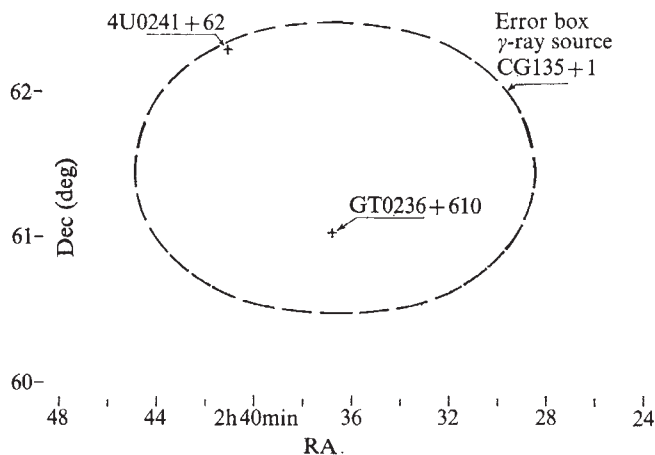


Fig. 2 The flux density of GT 0236+610 at a wavelength of 6 cm as a function of time.

reported⁸. The brightest optical object in the 30'' X-ray error circle has a redshift $z = 0.044$ and coincides within 2'' with a point like radio source⁸. They conclude that 4U0241+62 is either a QSO or Seyfert nucleus at a distance of ~ 260 Mpc. The results from the COS B satellite on the distribution of γ -ray sources indicate that γ -ray sources are galactic¹ which argues against an association of CG135+1 with 4U0241+62. At present, the only confirmed γ -ray source identifications are with the Crab and Vela pulsars. However, several groups⁹⁻¹¹ have reported γ -rays from the direction of Cyg X-3 modulated at the 4.8 h period observed in the X-ray and infrared regions. (Other attempts^{1,12} to detect γ -rays from Cyg X-3 have been unsuccessful.) The similarity in radio properties of Cyg X-3 and GT 0236+610 suggests GT 0236+610 is a promising radio counterpart to CG135+1.

Future steps towards a better understanding of this interesting source should include: (1) an accurate radio position determination and radio spectral information, (2) HI absorption measurements for a distance determination and (3) sensitive search for radio pulsar, weak X-ray source and binary star counterpart.

Fig. 3 The position of the variable radio source is plotted in relation to the γ -ray source CG135+1 and the nearest X-ray source 4U0241+62.



This research was supported by a NRC of Canada grant and scholarship. We thank the National Radio Astronomy staff for their assistance. The NRAO is operated by Associated Universities Inc. under contract with the NSF.

Note added in proof: On the basis of a recent accurate radio position measurement¹³ with the very large array we have been able to confirm the identification of GT 0236+610 with the OB+star LSI+61°303. Sanduleak¹⁴ has reported weak H α emission from the star, and Crampton and Hutchings¹⁵ have found very broad emission lines at H α and H β (total width ~ 900 km s⁻¹) with sharp central absorption lines superimposed on the emission features.

P. C. GREGORY
A. R. TAYLOR

Physics Department,
University of British Columbia, Vancouver,
British Columbia, Canada

Received 6 February; accepted 17 February 1978.

1. Hermesen, W. *et al.* *Nature* **269**, 494-495 (1977).
2. Gregory, P. C. *et al.* *Nature* **239**, 440-443 (1972).
3. Whelan, J. A. J. *et al.* *Mon. Not. R. astr. Soc.* **181**, 259 (1977).
4. Gibson, D. M., Viner, M. R. & Peterson, S. D. *Astrophys. J. Lett.* **200**, L143 (1975).
5. Hardrop, J., Rohlf, K., Slettebak, A. & Stock, J. *Luminous Stars in the Northern Milky Way 1* (Hamburger Sternwarte and Warner and Swasey Observatory, Hamburg-Bergdorf) (1959).
6. Drilling, J. S. *Astr. J.* **80**, 128 (1975).
7. Herr, R. B. *Astr. J.* **74**, 200 (1969).
8. Apparao, K. *et al.* *IAU Circ. No. 3150* (1977).
9. Galpev, A. M. *et al.* *Proc. 14th Int. Cosmic Ray Conf.* Munich **1**, 95 (1975).
10. Vladimirov, B. M., Neshpor, Yu. I., Stepanian, A. A. & Fomin, V. P. *Proc. 14th Int. Cosmic Ray Conf.* Munich **1**, 118 (1975).
11. Lamb, R. C., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Thompson, D. J. *Astrophys. J. Lett.* **212**, L63 (1977).
12. McKechnie, S. P., Mount, K. E. & Ramsden, D. *Astrophys. J. Lett.* **207**, L151 (1976).
13. Hjellming, R., Hogg, D., Hvatum, H., Gregory, P. C. & Taylor, A. R. *IAU Circ. No. 3180* (1978).
14. Sanduleak, N. *IAU Circ. No. 3170* (1978).
15. Crampton, D. & Hutchings, J. B. *IAU Circ. No. 3180* (1978).

Thermal continuum from accretion disks in quasars

ACCRETION onto black holes in galactic nuclei is a promising power source for quasars^{1,2}, but observational proof has been elusive. I argue here that supermassive accretion in disks may explain the flat optical continuum observed from some QSOs and Seyfert galaxies.

Matter spiralling into a black hole of mass $M_H = (10^8 M_\odot)$ M_8 produces a power³

$$L \approx (0.2c^2) \dot{M} = (10^{46.1} \text{ erg s}^{-1}) \dot{M}_0 \quad (1)$$

where $\dot{M}_0 = \dot{M}/1 M_\odot \text{ yr}^{-1}$. Combined with the Eddington limit,

$$L_{\text{Ed}} = (10^{46.1} \text{ erg s}^{-1}) M_8, \quad (2)$$

this suggests that a luminous QSO may be powered by accretion at a rate $\sim 1 M_\odot \text{ yr}^{-1}$ onto a black hole of mass $\sim 10^8 M_\odot$.

Power law, nonthermal continua are characteristic of QSOs over a wide range of frequencies. However, thermal continuum emission from the hypothetical accretion disk is also expected. The energy released per unit surface area^{4,5} is

$$F \approx 3GM_H\dot{M}/8\pi r^3 = (10^{19.5} \text{ erg s}^{-1} \text{ cm}^{-2})(M_8^{-2}\dot{M}_0)r_*^{-3}, \quad (3)$$

where $r_* = r/r_g$ and $r_g = GM_H/c^2$. If much of this is radiated thermally, the local effective temperature is

$$T = (F/\sigma)^{1/4} = (10^{5.9} \text{ K})(M_8^{-1/2}\dot{M}_0^{1/4})r_*^{-3/4}. \quad (4)$$

The local emission will be concentrated near

$$\nu_B = 2.8 kT/h = (10^{16.7} \text{ Hz})(M_8^{-1/2}\dot{M}_0^{1/4})r_*^{-3/4} \quad (5)$$

One may assume that equations (3) and (4) apply outside some

radius $r_{\min}$, within which most of the energy is released non-thermally. A simple integration over $r \geq r_{\min}$ gives the thermal continuum

$$L_{\nu}^{\text{th}} = (10^{29.8} \text{ erg s}^{-1} \text{ Hz}^{-1})(M_8 \dot{M}_0)^{2/3} v_{15}^{1/3}, \quad (6)$$

with an exponential cutoff above

$$v_{\max} = v_B(r_{\min}) = (10^{15.8} \text{ Hz})(M_8^{-1/2} \dot{M}_0^{1/4})(r_{\min*}/20)^{-3/4} \quad (7)$$

The H β luminosity⁶ of 3C273 corresponds to an ionising luminosity $L_i \approx 10^{46.3} \text{ erg s}^{-1}$ for $H \approx 100$. An assumed total luminosity $L \approx 10^{46.6} \text{ erg s}^{-1}$ requires $\dot{M}_0 \approx 3$ and $M_8 \gtrsim 3$. The infrared (IR) fluxes⁶⁻⁸ of 3C273 are consistent with $f_{\nu} = 10^{-25.2} v_{16}^{-1.2} (\text{erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1})$, a typical nonthermal spectrum. Let us suppose that this nonthermal continuum extends to higher frequencies, and compare it with the thermal continuum described by equations (6) and (7) (take $r_{\min*} = 20$, corresponding to roughly equal thermal and nonthermal luminosities). Figure 1 shows that the nonthermal continuum dominates in the IR and far ultraviolet (UV), but the thermal component dominates in the optical. This qualitative behaviour is observed for 3C273, as shown by Figs 2 and 3. For wavelengths $\lambda \gtrsim 10^4$ and $\lambda < 1,500 \text{ \AA}$ (refs 9-12), the continuum is consistent with a power law; but there is an excess in the optical.

I propose that this blue continuum component of 3C273 is thermal continuum emission from the inner part of an accretion

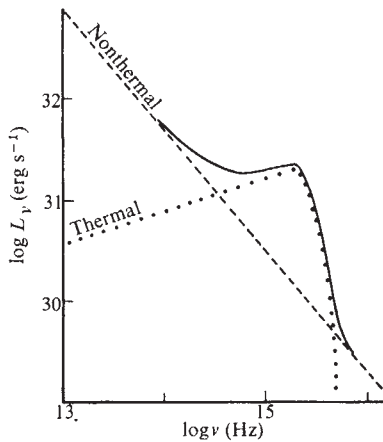


Fig. 1 The predicted thermal continuum for $M_8 = 10^{1.7}$ and $\dot{M}_0 = 0.7$. The solid line indicates the total continuum in this model.

disk that powers 3C273. This interpretation is supported by the general agreement between the observed thermal component and the predicted values of L_{ν}^{th} and $v_{\max}$. Furthermore, the excess continuum is inconsistent with Balmer or two-photon emission from the emission-line region⁶. Solving equation (6) with $\dot{M}_0 = 3$ for $L_{\nu} = 10^{31.0}$ at $v_{15} = 1$, we find $M_H \approx 10^9$. Accretion at $3 M_{\odot} \text{ yr}^{-1}$ can grow a hole of mass $10^9 M_{\odot}$ in less than the Hubble time. Spectral features will be severely broadened, as the circular velocity is $\sim 0.1c$ where the thermal continuum is emitted.

The present model suggests that many QSOs and Seyfert galaxies may show a power law continuum plus thermal emission from the disk at optical or UV frequencies. Indeed, many such objects do show a flattening or upturn of their continua in the blue. The QSO PHL957 ($z = 2.68$) shows a flat continuum and indications of a cutoff at the highest observed frequencies¹³. For $L_{\nu} = 10^{31.1}$ at 10^{15} Hz rest frequency and $v_{\max} \approx 10^{15.5} \text{ Hz}$, we find $M_H \approx 10^{9.2} M_{\odot}$ and $\dot{M} \approx 10 M_{\odot} \text{ yr}^{-1}$. The ratio $M_H/\dot{M}$ suggests that PHL 957 was luminous for much of the time before $z = 2.7$.

The above discussion does not specify the location of the nonthermal continuum source. But, the compact (10^{-3} arc s)

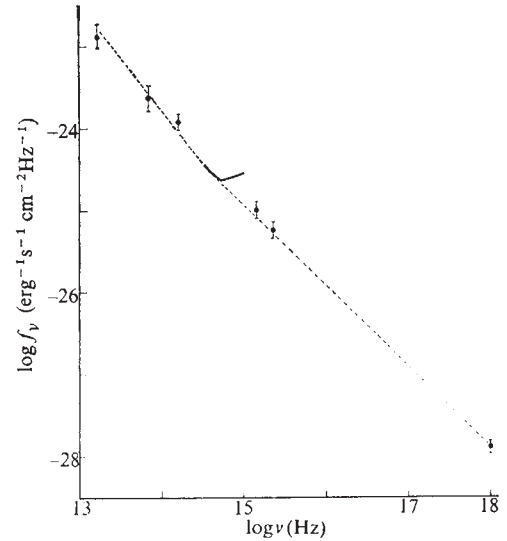


Fig. 2 Continuum observations of 3C273 at IR^{7,8}, optical⁶, UV¹⁰, and X-ray^{11,12} frequencies.

double radio sources in 3C273 and in some other quasars suggest that the disk emits narrow jets of energy in opposite directions along the rotation axis. These jets are a plausible source of the nonthermal continuum at IR, optical, UV, and X-ray frequencies. The jets might consist of a radiation-driven wind from the inner disk, with turbulence in the wind leading to *in situ* particle acceleration. The observed time-scales for optical continuum variability suggest that the non-thermal emission typically comes from the jets at heights $10^{17 \pm 2} \text{ cm}$, far above the plane of the disk.

Nonthermal ionising radiation striking the disk at $\sim 10^4 r_g$ can account for the broad emission lines¹⁴. The observed line profiles agree with computed profiles for plausible distributions of nonthermal emission along the beams. Radiation pressure compresses the gas in the broad line region to densities $\sim 10^9 \text{ cm}^{-3}$, in agreement with the absence of broad forbidden line emission. The resulting ratio of gas density to radiation flux agrees with photoionisation analyses of the emission-line intensities¹⁵.

Some QSOs^{16,17} show a fairly abrupt increase in continuum intensity at a frequency $\sim 10\%$ to the red of the Balmer limit, $v_{\text{Bac}} = 10^{14.92} \text{ Hz}$. Although this has been interpreted as Balmer continuum from the emission-line region, the feature begins further to the red of v_{Bac} than can be accounted for by the Doppler widths of the emission lines; and the energy in the feature is more than expected from the line intensities. I suggest that the feature in question is Balmer continuum from the inner accretion disk rather than the emission-line region. As the Balmer jump is in emission, the disk's upper atmosphere presumably has a temperature inversion, possibly resulting from heating by incident nonthermal continuum.

For example, let us consider QSO1217-02. The observed¹⁷ nonthermal continuum of QSO1217-02 indicates a total luminosity $\sim 10^{46} \text{ erg s}^{-1}$; considering equations (1) and (2), we take $\dot{M}_0 = 1$ and $M_8 = 10^{0.5}$. If the uncertain effect of the temperature inversion is neglected, equation (6) gives $L_{\nu}^{\text{th}} = 10^{30.1}$ at v_{Bac} . Thus, if the heated atmosphere of the disk is optically thin for $v > v_{\text{Bac}}$ and optically thick for $v \leq v_{\text{Bac}}$, the disk could add a component $\Delta L_{\nu} = 10^{30.1}$ at $v \geq v_{\text{Bac}}$. This agrees with the observed continuum excess $\Delta L_{\nu} \approx 10^{30.0}$ at v_{Bac} . From equation (4), we find $r_* \approx 10^2$ for the radius where $v_B(T) = v_{\text{Bac}}$. If this is the approximate radius where the presumed Balmer continuum originates, then Doppler broadening will redshift the onset of the feature an amount $\Delta v/v = r_*^{-1/2} \approx 10^{-1}$. This agrees with the observed wavelength of the upturn in the continuum of QSO1217-02. The

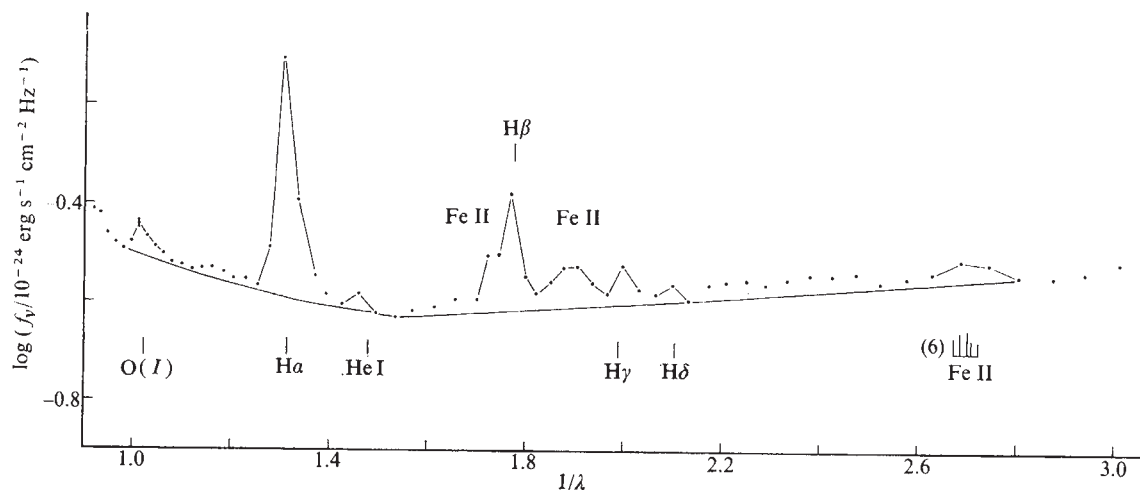


Fig. 3 Optical and near IR spectrum⁸ of 3C273. The change of slope at $\lambda^{-1} > 1.6 \mu\text{m}^{-1}$ clearly is a continuum feature beginning far to the red of the Balmer limit. (Reproduced with permission of *Astrophys. J.*).

absence of detectable line emission from the Balmer continuum zone is easily understood. The local surface brightness of the disk cannot exceed the Planck function, even in the centre of a line¹⁸. The luminosity in H β , from the zone in question, is therefore limited to $L(\text{H}\beta) \geq 4\pi B_\nu(T)(\pi r^2)\Delta\nu$, where $\Delta\nu$ is the line width of the local emission. For the above dimensions, $L(\text{H}\beta) \geq 10^{40} \Delta V_6 \text{ erg s}^{-1}$, where ΔV_6 is the turbulent velocity dispersion in the atmosphere in units 10^6 cm s^{-1} . Compared with the observed¹⁷ H β luminosity, $\sim 10^{43.6} \text{ erg s}^{-1}$, the H β luminosity from the Balmer continuum zone is negligible, even if there is highly supersonic turbulence in the disk.

I thank J. C. Wheeler, M. Rees and H. Netzer for useful discussions. This work was partly supported by NASA.

G. A. SHIELDS

Department of Astronomy,
University of Texas at Austin,
Austin, Texas 78712

Received 24 November 1977; accepted 14 February 1978.

1. Lynden-Bell, D. *Nature* 223, 690–694 (1969).
2. Salpeter, E. *Astrophys. J.* 140, 796–800 (1964).
3. Thorne, K. *Astrophys. J.* 191, 507–519 (1974).
4. Pringle, J. & Rees, M. *Astr. Astrophys.* 21, 1–9 (1972).
5. Shakura, N. & Sunyaev, R. *Astr. Astrophys.* 24, 337–355 (1973).
6. Oke, J. B. & Shields, G. *Astrophys. J.* 207, 713–724 (1976).
7. Rieke, G. & Low, F. *Astrophys. J. Lett.* 176, L95–L100 (1972).
8. Grasdalén, G. *Astrophys. J. Lett.* 208, L11–L12 (1976).
9. Davidsen, A., Hartig, G. & Fastie, W. *Nature* 269, 203–206 (1977).
10. Wu, C.-C. *Astrophys. J. Lett.* 217, L117–L120 (1977).
11. Kellogg, E. *et al. Astrophys. J. Lett.* 165, L49–L54 (1971).
12. Bowyer, C., Lampton, M., Mack, J. & de Mendonça, F. *Astrophys. J. Lett.* 161, L1–L8 (1970).
13. Lowrance, J., Morton, D., Zucchini, P., Oke, J. B. & Schmidt, M. *Astrophys. J.* 171, 233–252 (1972).
14. Shields, G. *Astrophys. Lett.* 18, 119–123 (1977).
15. Davidson, K. *Astrophys. J.* 218, 20–32 (1977).
16. Baldwin, J. *Astrophys. J.* 201, 26 (1975).
17. Baldwin, J. *Mon. Not. R. astr. Soc.* 178, 67P–74P (1977).
18. Rees, M. *Ann. N. Y. Acad. Sci.* (in the press).

Five-day oscillation in East African low-level jet

EASTERN edges of the continents are favourable for the deflection of tropical easterlies from the winter hemisphere to the summer hemisphere. Low-level southerly jet stream along the East African coast is the most pronounced manifestation of this deflected air current which feeds into the Indian monsoon current. We have subjected the jet stream data to a smoothing process to filter out noise and then found that this jet stream has a well marked 5-d periodicity.

We discuss here whether this 5-d period is related to the well-known 5-d pulsations of the Indian monsoon activity.

During the northern summer, the air in the northern-hemispheric tropics is continually warmed up while that in the southern-hemispheric tropics remains relatively cool. A cross-equatorial circulation is set up in the form of equatorial cell¹ with air from the Southern Hemisphere going to the Northern Hemisphere in the lower troposphere and in the opposite direction in the upper troposphere. There are preferred belts for this cross-equatorial flow, generally near the eastern edges of the continents which obstruct the easterly trades. After hitting the eastern edge of a continent, the tropical easterly flow must be deflected either northwards or southwards. During the northern summer season, there is a pressure gradient force from the winter hemisphere to the summer hemisphere. Furthermore, the easterlies hitting the continental barrier in the Southern Hemisphere already have a component from south to north (southeasterly trades). These factors combine to make the eastern edges of the continents favourable zones for cross-equatorial flow from Southern Hemisphere to Northern Hemisphere during the season in question. During the northern-hemispheric summer, the preferred belts for this cross-equatorial southerly flow are near north-east Brazil, East Africa and north-east Australia². The strongest flow is along the East African coast in the form of a low-level jet. This was discovered by Bunker³ and Findlater⁴ and has become the subject of several observational⁵ and theoretical^{6,7} investigations. Krishnamurti *et al.* simulated many features of this jet with a barotropic frictionless model by incorporating the East African and Malagasy Mountains, Earth's rotational β -effect and realistic lateral boundary conditions along long 75°E at 1.5 km level. Without mountains, the solution showed no resemblance to the jet, confirming the important part played by the physical barrier on the eastern edges of the continents in this cross-equatorial flow.

We have recently completed another observational study of this low-level jet. To eliminate observational errors, we smoothed each pilot balloon flight by fitting 5th degree polynomial to the computed meridional component of wind up to the 2.1 km level, at Garissa (lat 0.5°S, long 39.5°E, 0.14 km above sea level), and Mandera (4°N, 41.5°E, 0.28 km above mean sea level). We used 0001 GMT flights for 5 yr (1970–74) at Garissa and for 4 yr (1971–74) at Mandera during the 3-month period (June–August) when the jet has a peak intensity. The 92-d time-series at each of

the two stations in each year was subjected to spectral analysis for determining short period oscillations using the Jenkins and Watts⁸ method. The spectral estimates were tested, using the Tukey window for white noise at 95% confidence limit after the possibility of red noise was eliminated. In each year, in each station, we could clearly see the 4–5-d period oscillation as shown in Fig. 1.

Oscillations of this period in the tropical region have assumed great significance during the past few years. Standing oscillations of this period in Afro-Asian monsoon regions were first detected by Eliot⁹. Similar oscillations were detected by Frolow¹⁰ on the two sides of the Atlantic and in west central Pacific by Palmer and Ohmsted¹¹, 4–5-d migratory oscillations have also been recently discovered^{12,13}.

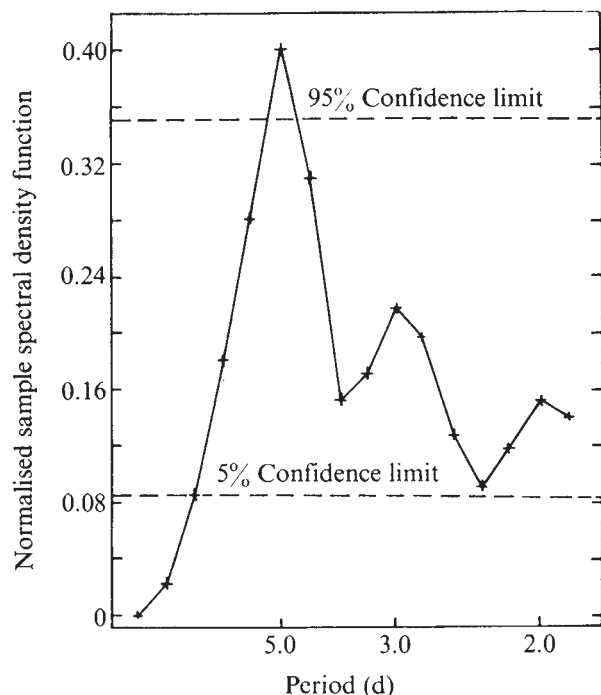


Fig. 1 Spectral analysis for determining short period oscillations at Garissa (1974) 1.3 km level using a Tukey window.

The 4–5-d oscillations have long been known to exist in Indian monsoon activity and these have been associated with 'surges' or 'pulses' from the Southern Hemisphere¹⁴. These pulsations were originally noticed by the strengthening of winds over the south-west Arabian Sea. But at that time, the existence of East African low level jet was not known.

We now believe that the 4–5-d oscillation which we have detected in the low level jet winds at Garissa and Mandera is perhaps associated with the 4–5-d oscillation in the Indian monsoon activity and with the 4–5-d oscillation first detected by Eliot⁹. We need cross-spectral analyses relating pulsations in the East African low level jet to other monsoon phenomena in south-east Asia. This work is in progress.

We also need analyses of other pilot balloon observations in Africa and then through co-spectral analysis to get to know more properties of this oscillation. However, the 'signal' is generally weak and gets lost in the 'noise', making the analysis of raw data difficult and often unproductive. We took a great deal of time and effort to smooth the data before subjecting them to spectral analysis. It was only after smoothing the daily wind data that we could clearly detect the 4–5-d oscillation in the East African low-level jet.

The oscillation was only faintly visible when we subjected the unsmoothed data to the spectral analysis.

T. NGARA
G. C. ASNANI

Department of Meteorology,
University of Nairobi, Kenya

Received 23 November 1977; accepted 17 February 1978.

1. Asnani, G. C. *Nature* **214**, 73–74 (1967).
2. Newell, R. E. *et al.* *The General Circulation of the Tropical Atmosphere and Interactions with Extra-tropical Latitudes 1* (MIT Press, Cambridge, Massachusetts, 1972).
3. Bunker, A. E. *Proc. Symp. Met. Results of the Int. Indian Ocean Exped.* 3–12 (1965).
4. Findlater, J. *Met. Mag.* **95**, 353–364 (1966).
5. Findlater, J. *Geophys. Mem.* No. 115, (HMSO, London, 1971).
6. Washington, W. & Daggupatty, S. *Mon. Weath. Rev.* **103**, 105–144 (1975).
7. Krishnamurti, T. N. *et al.* *J. Atmos. Sci.* **33**, 2350–2362 (1976).
8. Jenkins, G. M. & Watts, D. G. *Spectral Analysis and its applications* (Holden-Day, San Francisco, 1968).
9. Eliot, J. *Ind. met. Mem.* **6**, 89–160 (1895).
10. Frolow, S. *Bull. Am. met. Soc.* **23**, 239–254 (1942).
11. Palmer, C. E. & Ohmsted, W. D. *Tellus* **8**, 495–507 (1956).
12. Madden, R. & Julian, P. J. *atmos. Sci.* **29**, 1464–1469 (1972).
13. Misra, B. M. *Ind. J. met. Geophys.* **24**, 251–256 (1973).
14. Malurkar, S. L. *Ind. J. met. Geophys.* **3**, 303 (1953).

Angular localisation of proportional chamber avalanche

CHARPAK¹ points out in his review on multi-wire proportional chambers, that despite previous considerations² there is some evidence that the discharge in a proportional chamber does not necessarily spread uniformly round the wire. We have been examining in some detail avalanche localisation in proportional chamber operation during the past year or so, and give here a brief quantitative account of some of our findings. These were initially rather unexpected.

The most striking evidence of avalanche localisation is obtained from observations of the charges induced on neighbouring wires. These form particularly sensitive probes for examining the avalanche position, and chamber geometry and operation are not disturbed in any way. The situation is shown schematically, but not to scale, in Fig. 1, where A is the particular wire at which avalanche occurs and B is an adjacent wire at the same potential. A and B are connected to charge-sensitive amplifiers, of conversion gain $1/C$, followed by pulse shaping networks. A positive ion of charge q starting at time $t = 0$ at the surface of wire A, at angular position α , will drift away from A along the field line defined by α with velocity proportional to the local field value. The induced charges q_B and q_A may then be calculated by standard electrostatic theory as a function of the position of q and hence as a function of t . As in the creation of the positive ion charge q , an equal negative, electron charge $-q$ must have been deposited at $t = 0$ on wire A, the charge-sensitive amplifier output voltages are $-q_B/C$ and $(q_A - q)/C$. The pulse heights v_B and v_A from the pulse shaping networks may thus be calculated as a function of the starting angular position α . These calculations may be extended to describe the effect on v_B and v_A of an initial angular distribution of positive ion charge.

The important experimental observation, Fig. 2, is that the observed ratio $-v_B/v_A$ depends strongly upon the x -coordinate of a finely collimated soft X-ray beam, showing that the angular distribution of the positive ions round the wire is non-uniform and that the angular centroid of the distribution depends on the direction from which the initiating electrons have drifted. For normal chamber geometry the field lines near the cathode have equation $x = s(1 - 2\alpha/\pi)/2$ independent of y , where s is the anode wire spacing. Thus for absorption of a soft X-ray photon at the point X, coordinate x_0 in Fig. 1, the initial electrons drift along the field lines indicated and, ignoring diffusion, the angular centroid of the resultant avalanche at the wire A is $\alpha = \pi(1 - 2x_0/s)/2$.

Figure 2 shows experimental measurements of $-v_B/v_A$ plotted against x_0 , the x -coordinate of an X-ray beam of energy

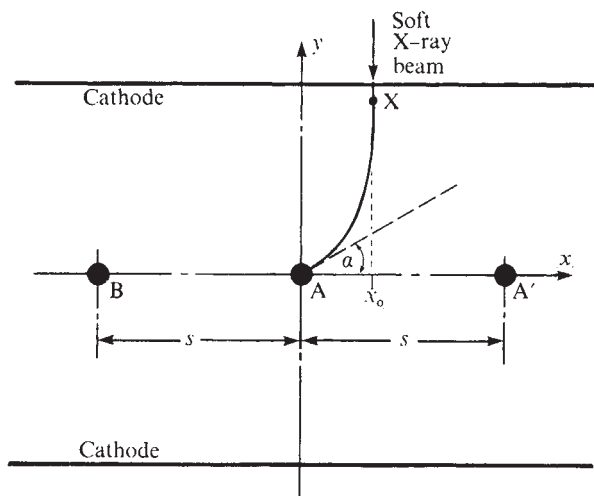


Fig. 1 Schematic diagram of conventional multi-wire proportional chamber. The angular position of the avalanche on wire A is obtained from measurement of the induced charge signal on wire B.

5.9 keV and collimated width 0.2 mm. The chamber cathode-anode separation was 11 mm, the anode wire separation, s , 5 mm and the diameter of the anode wires 25 μm . The chamber gas was argon/10% methane at atmospheric pressure. The pulse shaping networks consisted of passive, single RC integrating and differentiating circuits of equal time constant, 10 μs . The full curve is the result of calculations using the model described above, with zero angular spread. Also indicated on Fig. 2 are the theoretical predictions for rectangular positive ion distributions spreading round 50% and 100% of 2π (that is, a completely uniform spread). If x_0 exceeds $s/2$ then the avalanche

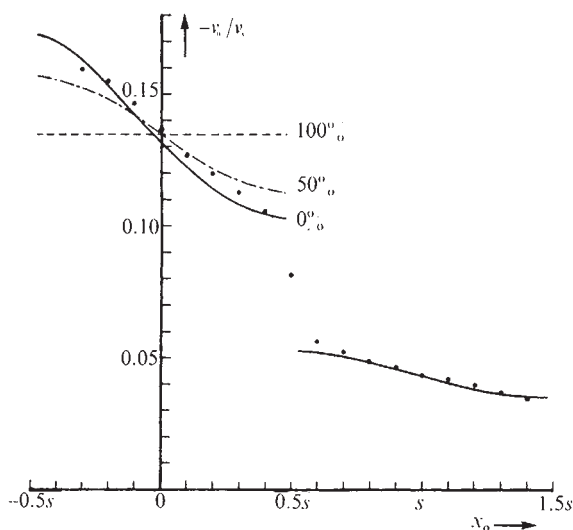


Fig. 2 Experimental measurements of the ratio of pulse heights $-v_B/v_A$, in the charge-sensitive pulse shaping systems connected to wires B and A, plotted against X-ray beam position x_0 . The curves are theoretical predictions for rectangular distributions spreading round 0%, 50% and 100% of the anode wire circumference.

switches over to the next wire A'. It is interesting that the wire B can still sense to some extent the avalanche position on A'. This is shown by the non-zero slope of the lower set of experimental results representing $-v_B/v_A$. The full curve is again the theoretical prediction, with zero angular spread.

Such measurements show the surprising degree to which angular localisation of the avalanche occurs, despite the diffusion which must occur during electron drift. Possible extensions of imaging MWPC operation are suggested as a result

of these investigations (such as positional interpolation between anode wires). A detailed report of this experimental and theoretical work and of similar investigations on cathode-induced signals is being prepared.

E. MATHIESON
T. J. HARRIS
G. C. SMITH

Physics Department,
The University,
Leicester, UK

Received 16 December 1977; accepted 1 February 1978.

1. Charpak, G. *Nature* **270**, 479-482 (1977).
2. Charpak, G., Rohm, D. & Steiner, H. *Nucl. Instrum. Meth.* **80**, 13-34 (1970).

The impossibility of comminuting small particles by compression

COMMINTION and crushing are important processes designed to break large particles into smaller ones by compression. Under such compression, cracks can propagate through a particle to split it into two or more fragments of smaller size and increase surface area if the material is sufficiently brittle. The appreciation of brittleness as applied to solids under compression is, therefore, basic to an understanding of comminution. This note analyses the brittleness of an ideally shaped particle subjected to compressive forces. New equations, based on the old ideas of Rittinger¹ and Griffith² are presented to explain the fracture of such a particle and these are verified by experiments on a model material, specially chosen to demonstrate the effects at the macro-level. The equations state explicitly that compressed brittle particles should increase in strength as they are made smaller, a fact well known from macroscopic fracture results. In addition, the equations predict a critical particle size, verified by experiment, below which crack propagation is impossible under compressive forces. Bodies smaller than this appear ductile and are squashed flat when attempts are made to comminute them by crushing.

The fact that brittle substances, such as rocks, seem stronger when made smaller was noted in the eighteenth century by Gauthey³ and has been confirmed using a variety of frangible materials⁴⁻⁸. This size effect in crushing remained a curiosity until it was demonstrated^{9,10} by applying the Griffith² energy criterion of fracture to the idealised geometry of Fig. 1a, that the crack propagation stress should be inversely proportional to the square root of sample size. In fact, the cracking force is

$$F = \frac{b}{(1-w/d)} \left(\frac{2ERd}{3} \right)^{1/2} \quad (1)$$

where E is Young's modulus; R , the fracture energy; w , d and b are dimensions, as shown in Fig. 1. It therefore requires a larger mean compressive stress to crack a smaller specimen.

Consider now how the particle shown in Fig. 1b will behave when compressed. The sharp edge will first be depressed plastically by the rigid platen to give a flat area bw

$$F = Ybw \quad (2)$$

where Y is the yield stress. Substituting this into equation (1) shows that the force needed for cracking is given by the solution to the quadratic

$$\frac{1}{Yd} \left(\frac{F}{b} \right)^2 - \frac{F}{b} + \left(\frac{2ERd}{3} \right)^{1/2} = 0 \quad (3)$$

For large particles this solution gives a cracking force dependent

on $d^{1/2}$ as before. But as the particle size is reduced to a critical value

$$d_{\text{crit}} = \frac{32ER}{3Y^2} \quad (4)$$

the cracking force rises rapidly. Below this size cracking is impossible and gross yielding must therefore occur.

Polystyrene (Shell Carinex) was chosen as the model material because its properties ($E = 2.8 \text{ GNm}^{-2}$, $Y = 80 \text{ MNm}^{-2}$, $R = 960 \text{ Nm}^{-1}$) on substitution in equation (4) led to a critical dimension of 4.48 mm, a convenient experimental value. There were additional advantages of specimen preparation and transparency.

The polymer was compression moulded into sheet, a crack was initiated in the sheet by pressing with a sharp knife and a specimen was then cut to include the crack. Various sizes of samples were produced with dimensions in the ratio $l/d = 2.5$. The thickness b was a constant 3 mm and the crack length $l/2$.

The compression results are shown in Fig. 2, where the failure force is plotted against the width d of the specimens. For large samples, cracking occurred at a force proportional to d^3 . However, as the specimen size approached 4 mm the cracking force started to rise in accordance with equation (3). Below this size the cracks did not propagate and the compression force corresponded to that required for gross yielding. The experimental transition from brittle to ductile behaviour was at a size of 3.6 mm as opposed to 4.48 mm from equation (4).

These ideas, that particles increase in compressive fracture strength as they are made smaller, and ultimately become ductile at a critical size, are thought to be relevant to comminution despite the simplifications involved in the analysis: the special particle shape, the neglect of crack initiation, the constraint of slow cracking into two pieces, and the ignoring of temperature and rate effects. For example, it is well known¹¹ that comminuting energy is utilised more effectively for large particles, reflecting the possibility that smaller fragments are deforming plastically rather than cracking, and thereby wasting much of the energy input.

Again it has long been recognised that there is a size limit to the crushing of materials. After a certain period of comminution the particles maintain the same size and surface area on average, and this phenomenon has been explained by a variety of mechanisms (see ref. 11, pp 87–91) including the removal of initiating flaws, aggregation of small fragments, and even chemical reaction. Equation (4) demonstrates that the theory of fracture can itself account for the comminution limit. For example, it predicts a crushing limit of around 1 μm for calcium carbonate. This value is to be compared with the average particle size of 0.8 μm for calcium carbonate after prolonged milling¹². Plastic yielding, as opposed to cracking, has been observed in compressed calcium carbonate particles about 4 μm in size¹³.

Fig. 1 a, Cracking of a brittle particle under compression; b, the shape of the experimental specimens.

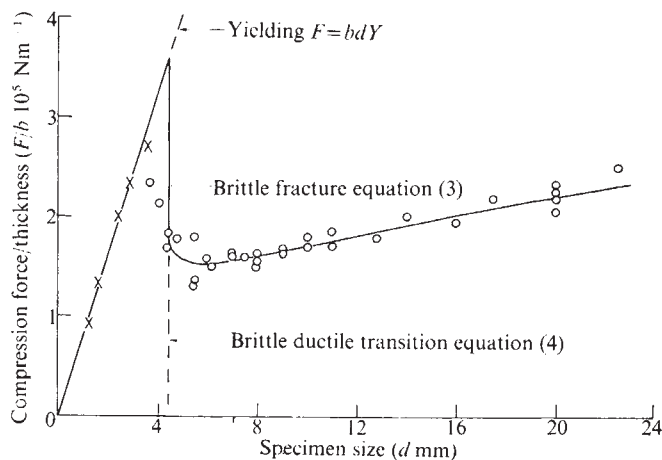
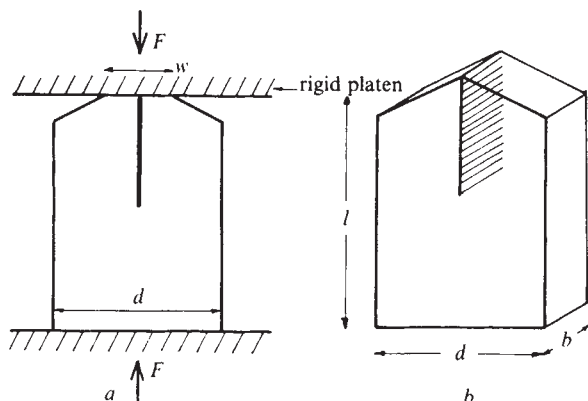


Fig. 2 Compression results for a range of specimen sizes showing cracking (○) of large samples and yielding (×) of small ones.

Of course, the plasticity of compressed small particles has been appreciated since the pioneering studies on coal by Boddy in 1943¹⁴. That this plasticity was associated with hardness, where hardness is roughly three times the yield stress¹⁵, was understood by Parish¹⁶. For coal, the transition from squashing to cracking occurred near a particle size of 5 μm (ref. 17). From equation (4) a transition in the micrometre range would be expected.

These size arguments are relevant not only to crushing but also to other processes connected with the brittleness and ductility of materials. For example, brittle substances may be indented plastically with a sharp tool, providing the indentation is below a certain size, above which cracking takes place. For calcium carbonate this critical size is 3 μm (ref. 18), a value comparable with that calculated from equation (4). Another example is that of cutting, which can only occur if the depth of cut is sufficiently small to prevent cracking. Ductile machining swarf has been observed when glass was cut by very fine tungsten carbide tools¹⁹. The depth of cut when this became apparent was 0.5 μm , as compared with a figure of 0.9 μm from equation (4).

In conclusion as a compressed brittle body is made smaller, its fracture strength apparently increases until, at a critical size of $32ER/3Y^2$, crack propagation becomes impossible. Particles below this size are ductile in compression. This brittle/ductile transition dimension seems to be relevant not only to the comminution process but also to indentation and cutting.

K. KENDALL

ICI Research and Development,
The Heath, Runcorn,
Cheshire, UK

Received 7 November 1977; accepted 13 February 1978.

1. Von Rittinger, P. R. *Lehrbuch der Aufbereitungskunde* (Ernst und Korn, Berlin, 1867).
2. Griffith, A. A. *Phil. Trans. R. Soc. A* **221**, 163 (1920).
3. Timoshenko, S. *History of strength of materials*, 58 (McGraw-Hill, London, 1953).
4. Rice, G. S. & Enzian, C. *US Bur. Mines Bull.* **303**, 5 (1929).
5. Gaddy, E. L. *Virginia Poly Tech. Inst. Eng. Extension St. Ser.* **112** **44**, 10 (1956).
6. Jellinek, H. H. G. *Proc. phys. Soc. Lond.* **71**, 797 (1958).
7. Holland, C. T. *Min. Congr. J.* **48**, 325 (1962); **48**, 466 (1962).
8. Evans, I. & Pomeroy, C. D. *Strength Fracture and Workability of Coal* ch. 3 and 4 (London, Pergamon, 1966).
9. Kendall, K. *J. mater. Sci.* **11**, 1267 (1976).
10. Kendall, K. *Proc. R. Soc. (in the press)*.
11. Lowrison, G. C. *Crushing and grinding* ch. 1–5 (Butterworths, London 1974).
12. Gregg, S. J., *Chem. Ind.* **611** (1968).
13. Steier, V. K. & Schonert, K. *Dechema Monog.* 69 Pt 1, 167 (Dechema, Frankfurt, 1972).
14. Boddy, R. G. H. B. *Proc. Conf. Ultrafine Structure of Coals and Cokes*, 336 (BCURA, London, 1943).
15. Tabor, D. *Br. J. appl. Phys.* **7**, 159 (1956).
16. Parish, B. M. *Br. J. appl. Phys.* **18**, 233 (1967).
17. Millard, D. T., Newmann, P. E. & Phillips, T. W. *Proc. phys. Soc.* **B68**, 723 (1955).
18. Shaw, M. C. *Proc. natn. Acad. Sci. U.S.A.* **40**, 394 (1954).
19. McClintock, F. A. & Argon, A. S. *Mechanical Behaviour of Materials* 495 (Addison Wesley, Reading USA, 1966).

Interfacial noble-metal corrosion in metal to ceramic reaction welding

WELDING refractory ceramic oxides to a range of transition metals in oxidising atmospheres may involve two apparently distinct bonding mechanisms^{1,2}. In type 1 bonds, which involve the noble metals, a sharp phase discontinuity at the metal-ceramic interface is maintained, even after prolonged annealing in oxygen. No intermediate oxide phase, other than the original ceramic, is observed. The substantial bond strengths³ reach their maximum value after a few hours and do not deteriorate with time. Residual thermal stresses are small as shown by high resolution hot-stage electron microscopy⁴. In contrast, in type 2 bonds, in which the more reactive transition metals are used, an intermediate (quaternary) oxide layer is clearly visible, its thickness growing with time. The bond strength depends on the corrosion behaviour of the metal and the mechanical strength of the intermediate oxide layer, as metal cations diffuse by bulk and grain boundaries into the ceramic. It is influenced by the duration of the anneal, as the oxidation process continues. Here we report the experimental observation that the couples Pd/MgO and Pd/ Al_2O_3 form type 2 bonds below 800 °C and strong type 1 bonds above that temperature, when annealed in air.

Figure 1 shows two Pd/MgO couples. Squares of polished {100} faces of single crystal MgO allow us to view the metal-ceramic interfaces directly. The characteristic metallic appearance of the type 1 bond (a) and the heavily oxidised type 2 interface (b) are clearly visible. This suggests that the two bonding types differ only in extent, and not in principle. Therefore, the creation of an interfacial oxide layer originating from the metal component, and structurally compatible with both metal and ceramic oxide, is common to both bonding mechanisms. Above 800 °C PdO should dissociate under 1 atm O_2 pressure. However, the near surface region at a palladium-ceramic interface is a high energy state relative to the bulk, and oxidised species, unstable in the bulk, can exist at this interface.

The thickness of some such interfacial noble metal oxide layers can be estimated approximately from their known free energies of formation (Fig. 2) and the interfacial energies of metal-ceramic couples (Fig. 3). For the metals in Fig. 2 the process of reaction welding to a refractory oxide, such as alumina, takes place above 700 °C. A metal whose oxide has a negative free energy of formation will

Fig. 1 Two bonding mechanisms for Pd/MgO couples above and below 800 °C. Type 1 couple (a) was heated for 13 h at 1,400 °C, the type 2 couple (b) for 250 h at 800 °C; both in air. The metal-ceramic interfaces are viewed through polished {100} faces of the MgO single crystals.

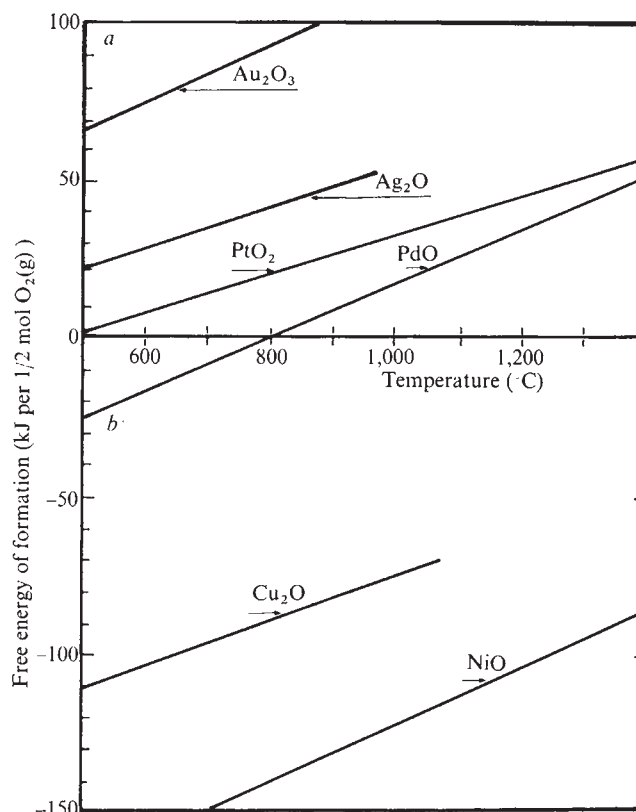
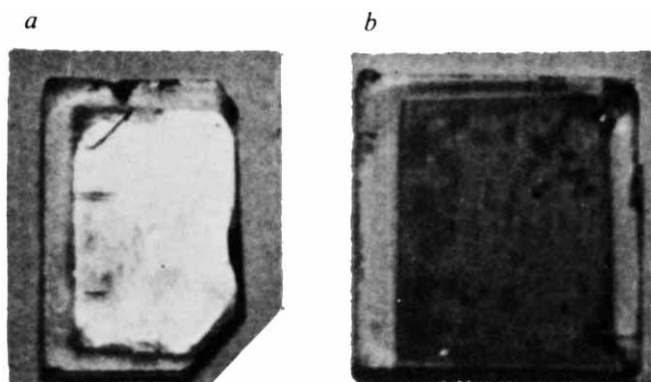


Fig 2 Free energies of formation for metal oxides^{13,14}: a, type 1; b, type 2.

form a type 2 bond (Cu, Ni) in contrast to the noble metals Ag, Au and Pt which form type 1 bonds.

Figure 3 is the decrease in interfacial free energies (in kJ mol^{-1}) from the junction towards the bulk of the ceramic for interfaces between relevant metals and $\alpha\text{-Al}_2\text{O}_3$, calculated from data by Hondros⁵ and McClean and Hondros⁶. This diagram suggests that the free energies of formation of noble metal oxides in the surface layers should not be calculated using the zero point energy of the bulk, but a significantly higher reference level, raised by the local value of the interfacial energy for the particular system.

Comparison of Figs 2 and 3 gives an order of magnitude estimate of the thickness of an interfacial noble metal oxide layer in Al_2O_3 , assuming this oxide to be of the simple secondary type. This shows that the noble metal oxide can be little more than one or two lattice spacings deep. In that sense they are two dimensional surface compounds, possibly grossly non-stoichiometric and even amorphous, conditions which make their identification extremely difficult. They are likely to be further stabilised because of formation of quaternary oxides⁷.

The seemingly different behaviour of Pd/ Al_2O_3 and Pd/MgO couples above and below 800 °C suggests a new corrosion phenomenon for noble metals in the presence of ceramic oxides. Such electron transfer reactions can be studied by impedance dispersion techniques at high temperatures. In fact, several workers^{8,9} have postulated the existence of noble metal oxides at the electrode/electrolyte interface in defiance of classical thermodynamic considerations.

The presence of surface compounds, as proposed here, contrasts to the generally assumed absence of chemical reactions at noble metal-ceramic interfaces. It has an important bearing on the behaviour of metal-insulator and

metal-semiconductor interfaces, particularly in thin film technology. Rapid transport of ionic metal species, for example, through the interfacial region can explain the anomalous diffusion behaviour in epitaxial Au films on oxides¹⁰, or electro-mass transport in integrated circuitry¹¹, and the instability of non-equilibrated selective surfaces on photovoltaic and other semiconductors¹².

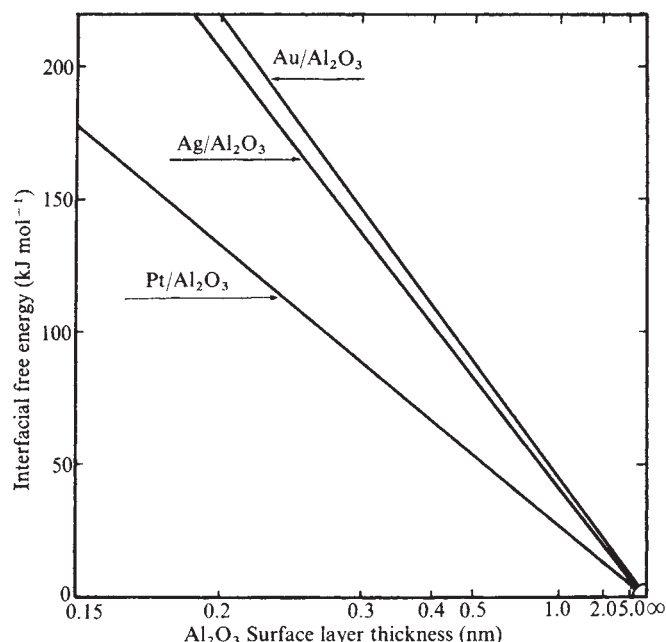


Fig. 3 The decrease of surface free energies in metal- Al_2O_3 interfaces^{5,6}, as a function of the thickness of the interfacial layer.

I thank Dr S. P. S. Badwal for assistance with the experiments and for helpful discussions with him and Dr N. J. Clark. Financial support from the Australian RGC is acknowledged.

H. J. DE BRUIN

School of Physical Sciences,
The Flinders University of South Australia,
Bedford Park, South Australia 5042,
Australia

Received 16 January; accepted 15 February 1978.

1. de Bruin, H. J., Moodie, A. F. & Warble, C. E. *J. Mater. Sci.* **7**, 909-918 (1972).
2. de Bruin, H. J., Moodie, A. F. & Warble, C. E. *J. Aust. Ceram. Soc.* **7**, 57-58 (1971).
3. Bailey, F. P. & Black, K. T. *J. Mater. Sci.* (in the press).
4. Lugton, F. D. & Warble, C. E. *Rev. Scient. Instrum.* **41**, 1793-1797 (1970).
5. Hondros, E. D. in *Interfaces Conf., Melbourne 1969* (ed. Gifkins, R. C.) 77-100 (The Australian Institute of Metals and Butterworths, Sydney, 1969).
6. McClean, M. & Hondros, E. D. *J. Mater. Sci.* **6**, 19-24 (1971).
7. Shannon, R. D., Rogers, D. B. & Prewitt, C. T. *Inorg. Chem.* **10**, 713-718 (1971).
8. Schouler, E., Kleitz, M. & Deportes, C. *J. chim. Phys.* **70**, 923-935 (1973).
9. Perfil'ev, M. V. & Pal'guev, S. F. in *Electrochemistry of Molten and Solid Electrolytes* (ed. Baraboshkin, A. N.) **4**, 153-157 (Consultants Bureau, New York, 1967).
10. Gupta, D. *Thin Solid Films* **25**, 231-244 (1975).
11. d'Heurle, F. M. & Rosenberg, R. *Phys. Thin Films* **7**, 257-310 (1975).
12. Harris, L. A., Gerstner, M. E. & Wilson, R. H. *General Electric Technical Information Series No. 77 CRD103* (1977).
13. Kubaschewski, O., Evans, E. D. & Alcock, C. B. *Metallurgical Thermochemistry* 4th edn (Pergamon, Oxford, 1967).
14. Steele, B. C. H. in *Electromotive Force Measurements in High Temperature Systems* (ed. Alcock, C. B.) (Institution of Mining and Metallurgy, London, 1968).

Anion deficiency in strontium titanate

FRANCO and REGI¹ diffraction patterns and lattice images of anion-deficient strontium titanate, SrTiO_{3-x} , show a sixfold superlattice along [111] and associated lattice images with irregular fringe spacings. The possibility of anion vacancy ordering was induced with reference to the systems (Ba,Sr)

MnO_{3-x} (ref. 2) and BaFeO_{3-x} (ref. 3). These results contrasted with earlier data on SrTiO_{3-x} ($0 \leq x \leq 0.5$) which indicated^{4,5} on the basis of powder X-ray diffraction, a simple cubic cell with virtually unchanging lattice constant over the whole range of x . Here we draw attention to two points concerning the nonstoichiometry of strontium titanate. First, unlike the manganate and ferrate of refs 2 and 3, SrTiO_3 is a simple cubic perovskite with cubic stacking of SrO_3 layers. Second, in contrast to the impression given, oxygen vacancy ordering had not been convincingly established at all until recently in any anion-deficient cubic-stacked perovskite. Examples of such nonstoichiometric systems are $\text{SrTiO}_{2.5-3.0}$ (refs 4, 5), $\text{SrVO}_{2.5-3.0}$ (ref. 4), $\text{SrFeO}_{2.78-3.0}$ (ref. 6) and $\text{Sr}_2\text{LiMoO}_{6.5}$ (ref. 7). I started to study the three strontium-transition metal oxide systems in the expectation that application of the more powerful diffraction techniques such as electron diffraction would reveal anion ordering processes possibly overlooked by X rays. This proved to be the case for the SrFeO_{3-x} system where an ordered phase, $\text{SrFeO}_{2.75}$, with superlattice reflections clearly apparent by electron diffraction, but not by powder X-ray diffraction, was found⁸.

Only a partial investigation of the titanium and vanadium systems was undertaken but these provided somewhat different results to those presented by Franco and Regi¹ and pointed to particular problems in sample preparation and characterisation which should be given some prominence in work of this type. A sample of nominal composition $\text{SrTiO}_{2.5}$ was prepared by firing a mixture of SrO, Ti metal powder (analysed by oxidation as $\text{TiO}_{0.02}$), and TiO_2 in a tantalum crucible inside an evacuated quartz tube at 1,200 °C for 2 d. The quartz tube was held inside an evacuated mullite tube to prevent oxygen ingress, which can be significant at 1,200 °C. The SrO was prepared by vacuum decomposition of SrCO_3 (all starting materials were Johnson-Matthey 'Specpure'), and all handling was done in an argon-filled dry box. The product, unlike that given in ref. 1, was a homogeneous dark-blue colour; it was cubic with no extra lines and a lattice constant of $3.9063(2)\text{\AA}$, a little higher than in refs 1, 4 and 5 [for SrTiO_3 prepared from SrCO_3 and TiO_2 we found $a_0 = 3.9047(1)\text{\AA}$].

Nevertheless, gravimetric analysis of the sample by oxidation showed the composition to be $\text{SrTiO}_{2.81}$, not $\text{SrTiO}_{2.5}$. The silica tube had completely devitrified and, although there was no reaction of the titanate with the tantalum, oxygen pick-up seems almost inevitable by this preparative route. Electron diffraction patterns did not show any convincing evidence for superlattice formation, see, for example, Fig. 1.

Another potential problem of a long reaction at 1,200 °C in a silica tube as performed by Franco and Regi¹ is pick-up of SiO_2 itself into the sample—note, for example, the first preparation of the phosphosilicate $\text{VO}(\text{P}_2\text{SiO}_8)$ (ref. 9) which could be responsible for colour variations through the sample. Our attempts at heating samples in an open tantalum crucible under vacuum or argon in a tantalum resistance furnace as in ref. 5 encountered the other particular preparative problem with this system, which can also occur in quartz, that of strontium-metal evaporation, and we found this route totally unsatisfactory. Thus, both the preparation and the characterisation of $\text{SrTiO}_{2.5}$ remain interesting problems still to be satisfactorily resolved, although $\text{SrTiO}_{2.8}$ does not seem to show vacancy ordering. I suggest that the only reliable preparation will be in a sealed, evacuated tantalum or similar tube, using a mixture such as $\text{Sr}_2\text{TiO}_4 + \text{TiO}$ as a preferred starting material to avoid the problems of handling SrO. The vacancy ordering, if observed, will surely not be related to the crystallographic shear of the ReO_3 -type oxides because of the blocking Sr^{2+} ion in the crystal and because of the absence of superlattice reflections on the powder patterns; any significant metal displacements such as occur in shear phases will produce readily visible superlattice lines, especially at such massive defect concentrations. The apparent absence of these tells us that, as in $\text{SrFeO}_{2.75}$, any ordering must involve primarily the oxygen atom, the weakest X-ray scatterer.

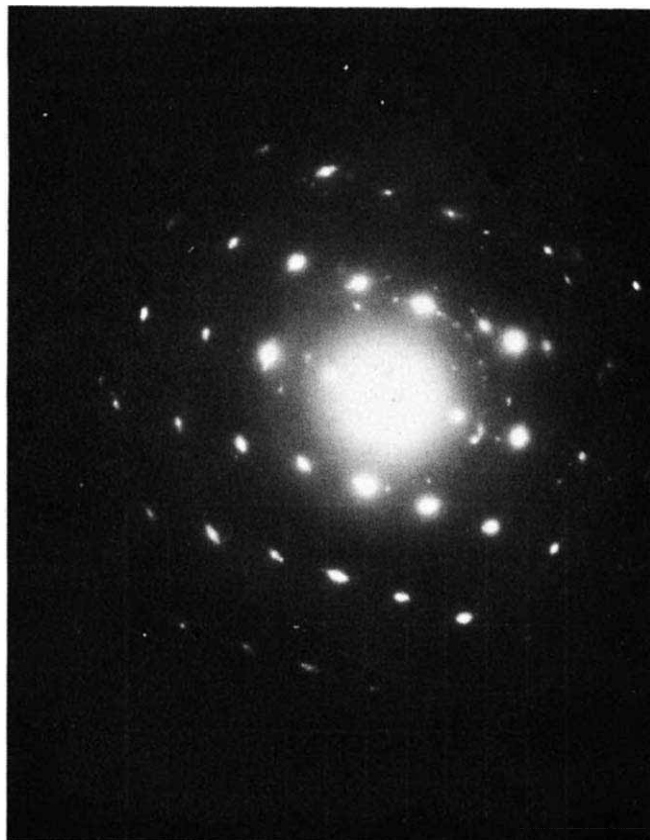


Fig. 1 Electron diffraction pattern of $\text{SrTiO}_{2.8}$ showing the (110) reciprocal lattice section. The pattern is heavily exposed and although there is slight contamination from overlapping patterns from associated crystals, no superlattice reflections are visible. (Compare Fig. 1b of ref. 8 for $\text{SrFeO}_{2.75}$ and Fig. 3 of ref. 1, which show the same sections.) Electron-diffraction patterns for other reciprocal-lattice sections of $\text{SrTiO}_{2.8}$ consistently revealed only the primitive perovskite reflections.

The study of SrVO_{3-x} by contrast gave some indications that at $x \sim 0.2$ some ordering effects were indeed apparent. $\text{SrVO}_{3.0}$ was cubic, however, with no extra lines, $a = 3.8415(2) \text{ \AA}$, slightly lower than found previously ($3.8424 \text{ \AA}$ and $3.8426 \text{ \AA}$; refs 10 and 11 respectively). The preparation was rather simple, not needing to handle SrO ; $\text{SrVO}_{2.85}$ was formed from reaction between SrCO_3 and V_2O_5 in hydrogen; a portion was oxidised to $\text{Sr}_2\text{V}_2\text{O}_7$, and the appropriate quantities of $\text{SrVO}_{2.87}$ and $\text{Sr}_2\text{V}_2\text{O}_7$ were heated in a quartz tube at $1,000^\circ\text{C}$.

B. C. TOFIELD

Materials Physics Division,
AERE, Harwell,
Oxfordshire, UK

Received 6 February; accepted 17 February 1978.

1. Franco, M. A. Alario & Regi, M. V. *Nature* **270**, 706–707 (1977).
2. Jacobson, A. J. & Horroxx, A. J. W. *Acta crystallogr.* **B32**, 1003–1008 (1976).
3. Jacobson, A. J. *Acta crystallogr.* **B32**, 1087–1090 (1976).
4. Kestigian, M., Dickinson, J. G. & Ward, R. J. *Am. chem. Soc.* **79**, 5598–5601 (1957).
5. McCarthy, G. J., White, W. B. & Roy, R. J. *Am. ceram. Soc.* **52**, 463–467 (1969).
6. MacChesney, J. B., Sherwood, R. C. & Potter, J. F. *J. chem. Phys.* **43**, 1907–1913 (1965).
7. Sabatier, R., Wathlé, M., Besse, J. P. & Baud, G. *J. inorg. nucl. Chem.* **33**, 1597–1613 (1971).
8. Tofield, B. C., Greaves, C. & Fender, B. E. *Mat. Res. Bull.* **10**, 737–746 (1975).
9. Tofield, B. C., Crane, G. R., Bridenbaugh, P. M. & Sherwood, R. C. *Nature* **253**, 722–723 (1975).
10. Chamberland, B. L. & Danielson, P. S. *J. Solid St. Chem.* **3**, 243–247 (1971).
11. Douglar, P., Fan, J. C. C. & Goodenough, J. B. *J. Solid St. Chem.* **14**, 247–259 (1975).

Single crystal analysis of the structure of stishovite

STISHOVITE is the highest pressure polymorph of silica. It possesses the rutile structure with a density of 4.28 g cm^{-3} and is characterised by sixfold coordination of silicon by oxygen. Previous structural studies^{1–3} of stishovite have been based on powder diffraction data. However, complete definition of its crystal structure requires single crystal studies. We have, therefore, undertaken a programme to synthesise single crystals of high pressure phases. We report here the successful synthesis of single crystals of stishovite up to $230 \mu\text{m}$ in length, and describe its crystal structure.

The high density of stishovite is a consequence of its octahedral coordination of silicon, in contrast to most known silicates in which the silicon atoms are tetrahedrally coordinated. Stishovite was first synthesised⁴ at high pressure, above 100 kbar, and soon after, it was discovered⁵ naturally in an impact breccia from Meteor Crater, Arizona, where it has been formed under the extremely high shock pressures generated by the fall of the Canon Diablo meteorite. Its crystal chemistry is of considerable importance in connection with interpretations of the structure of minerals which occur at great depths ($>600 \text{ km}$) in the Earth's mantle. Under the very high pressure existing in this region, the common minerals of the upper mantle transform to assemblages of very dense high pressure polymorphs characterised also by sixfold coordination of silicon^{6,7}.

The sample of stishovite studied was crystallised using a newly developed large Bridgman anvil type apparatus, equipped with an internal heating system. The starting material was A.R. grade Mallinckrodt silicic acid to which 40% of water was added as a flux. These components were sealed in a platinum capsule and held at approximately 90 kbar, and a temperature of 700°C , for 45 min, after which the charge was quenched and cooled to ambient temperature and pressure was released. When the capsule was opened many clear euhedral crystals of the rutile polymorph were obtained, ranging in size from 10 to $230 \mu\text{m}$.

Precession and Weissenberg photographs confirmed the space group to be $P4_2/m2_1/m2/m$. The intensity data were recorded on a computer controlled Supper equi-inclination diffractometer by the ω -scan method. The unit cell dimensions $a = 4.1772(7)$ and $c = 2.6651(4) \text{ \AA}$ were determined

Table 1 Atomic parameters of stishovite

Atom	No. per cell	x	y ($\times 10^4$)	z	β_{11} ($= \beta_{22}$) ($\times 10^4$)	β_{33}	β_{12}	B_{iso} ($\times 10^3$)
Si	2	0	0	0	23 (1)	42 (2)	2 (1)	152 (4)
O	4	3062 (1)	3062 (1)	0	36 (1)	65 (2)	–10 (1)	234 (5)

Estimated standard deviations are given in parentheses.

from a number of very high angle reflections. $\text{MoK}\alpha$ radiation reflected from a graphite crystal monochromator was used. Of the 238 independent reflections obtained, 236 were used in the final refinement—the 002 and 110 were thought to be affected by extinction.

The last cycle of refinement using anisotropic temperature factors gave an unweighted disagreement value, $R = 0.0153$. The final parameters are given in Table 1. The positional parameters are in good agreement with those reported from powder data³. Anisotropic temperature factors for both cation and anion, however, have significantly different values from those reported previously³. Bond distances and angles are shown in Table 2.

Table 2 Bond lengths and bond angles

		No. per cell
Si-O	1.8088 (1)	× 2
Si-O	1.7567 (1)	× 4
Si-O _m	1.7741	
O-O	2.5216 (2)	× 8
O-O	2.2891 (2)	× 2
Si-Si	3.2404	× 8
O-Si-O	81.34	× 2
O-Si-O	180.00	× 3
O-Si-O	90.00	× 8
O-Si-O	98.66	× 2

Estimated standard deviations are given in parentheses.

In this structure each silicon atom is surrounded by six oxygen atoms at the corners of a distorted octahedron. The octahedra share edges, forming bands running parallel to the *z* axis. The four coplanar bonds located in the (110) plane are 0.052 Å shorter than the two axial bonds lying normal to the (110) plane. The new values for the temperature factors show that the thermal motion of the oxygens is largest in a direction perpendicular to the line joining their centres. This can be explained by the extremely short O-O bond distance of 2.29 Å ($r_{\text{ox}} = 1.40$ Å) (ref. 8). The isotropic thermal parameters (*B*s) are considerably smaller than usual values for the tetrahedrally coordinated arrangement—a consequence of the close packing of the atoms.

This successful synthesis of crystals of stishovite demonstrates the feasibility of synthesising many other single crystals similar to those occurring in the Earth's mantle. We have already synthesised two other phases in this category and plan to extend this programme.

We thank Dr I. Threadgold for the use of the diffractometer for data collection.

W. SINCLAIR
A. E. RINGWOOD

Research School of Earth Sciences,
Australian National University,
Canberra, A.C.T., 2600, Australia

Received 16 January; accepted 3 March 1978.

1. Stishov, S. M. & Belov, N. V. *Dokl. Akad. Nauk. SSR* 143, 951 (1962).
2. Preisinger, A. *Naturwissenschaften* 49, 345 (1962).
3. Baur, W. H. & Khan, A. A. *Acta crystallogr.* B27, 2133–2138 (1971).
4. Stishov, S. M. & Popova, S. V. *Geokhimiya* 10, 837–839 (1961).
5. Chao, E. C. T., Fahey, J. J., Littler, J. & Milton, E. J. *J. geophys. Res.* 67, 419–421 (1962).
6. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
7. Liu, L. *Geophys. J. R. astr. Soc.* 48, 63–62 (1977); *Earth planet. Sci. Lett.* 36, 237–245 (1977); *Geophys. Res. Lett.* 4, 183–185 (1977).
8. Shannon, R. D. & Prewitt, C. T. *Acta crystallogr.* B25, 925–946 (1969).

Coesite and stishovite in the Vredefort Dome, South Africa

THE high pressure silica polymorphs, coesite and stishovite have been discovered in very thin pseudotachylite veins developed in Witwatersrand quartzites of the collar of the Vredefort Dome. They are associated with microcrystalline fibrous quartz which is interpreted as reverted diaplectic silica glass. The presence of these minerals strongly supports the interpretation of the Vredefort Dome as an astrobleme.

The dome of Vredefort is situated 100 km south-west of Johannesburg on the margin of the Potchefstroom basin and represents a well known ring structure 80 km in diameter. It consists of an Archaean granite core surrounded by concentric layers of sedimentary rocks which are generally overturned¹. Since 1936, several authors^{2–5} have suggested that this structure result from a large, deeply eroded astrobleme. This assumption was based on the unusual nature of the structure,

the presence of pseudotachylites, shatter cones, and deformation lamellae in quartz. The age of the event was established at 1,970 Myr from the dating of granophyre dykes which have been emplaced after the formation of the Dome⁶ and are assumed to represent impact melts. Nevertheless many objections to the impact theory have been expressed and the theory has met with only partial acceptance^{6,7}.

Eight samples of pseudotachylite were collected in the quartzites of the Witwatersrand Supergroup and tested for coesite and stishovite by Stöffler's method⁸. The concentration process is based on the fact that in hydrofluoric acid, quartz is more soluble than coesite, which itself is more soluble than stishovite. The material was powderised and partially attacked by a mixture of hydrofluoric and hydrochloric acids. The process was halted at regular intervals at which the insoluble residue was X-rayed and then treated again by the acids. The operation was repeated until only an insoluble residue remained. This proved to be predominantly rutile. Sample 1 collected in the Kimberley-Elsburg Quartzite on the farm Weltevrede 167 (27°38'54"E/26°50'51"S), contained coesite and stishovite, and sample 2 from the top of the Main Bird Quartzite on Ospan 238 (27°38'33"E/26°50'48"S) has only coesite. The other samples collected on Mara 395 and Vlakspuit 1180 (Hospital Hill Quartzite), as well as on Nooitgedacht 508 IQ (Kimberley-Elsburg Quartzite) were barren. In samples 1 and 2, the two most intense reflections of coesite are clearly visible in X-ray diffraction patterns of raw material, indicating a content of at least several per cent in each. By treatment, a concentrate containing about 90% coesite was obtained (Table 1). Stishovite, as in other natural occurrences, is much less abundant and it was not possible to obtain more than seven diffraction lines in the richest concentrates (Table 2). Its abundance is probably less than 0.1%.

Table 1 X-ray diffraction data on coesite

Sample 1*		ASTM 14-654	
<i>d</i> (Å)	<i>I</i> / <i>I</i> ₁	<i>d</i> (Å)	<i>I</i> / <i>I</i> ₁
6.16	6	6.20	8
4.37	4	4.38	4
3.42	30	3.43	30
3.09	100	3.09	100
2.75	15	2.76	8
2.69	15	2.69	10
2.33	6	2.33	4
2.29	7	2.29	8
2.18	7	2.18	4
2.03	4	2.03	6
1.83	3	1.84	4
1.78	8	1.79	8
1.71	11	1.71	12
1.70	9	1.70	10
1.65	4	1.66	2
1.58	6	1.58	6
1.541	7	1.545	10
1.497	2	1.501	4
1.416	1	1.418	2
1.405	3	1.409	4
1.341	13	1.345	12

*Cobalt radiation, quartz of the matrix as standard.

These two samples were collected in poorly developed pseudotachylite. For instance at the site of sample 1, pseudotachylite forms an isolated vein ~0.5 mm thick extending for <1 m. Under the microscope it is seen to consist of a very finely crystallised mixture of coesite grains in a quartz matrix. The country rock is a coarse-grained quartzite with mostly large monocrystalline quartz grains in a sericite matrix. In the immediate vicinity of the pseudotachylite, the large quartz crystals are altered to a minutely granular and radiating quartz identical to flint or chert (Fig. 1). This material is interpreted as reverted diaplectic quartz, as it is logical to assume that any amorphous silica, whether of sedimentary origin or

the result of shock metamorphism, can converge to a similar texture by recrystallisation. In this presumed reverted diaplectic quartz, coesite forms granular crystals and rosettes of needles reaching to a maximum length of 100 μm (Fig. 1). This often shows resorption to quartz (Fig. 1). Stishovite develops in a similar manner to coesite (Fig. 1), except that it is less abundant and occurs as shorter needles ($\leq 10 \mu\text{m}$). Both coesite and stishovite were examined by electron microprobe and found to be pure silica.

Coesite, stishovite and diaplectic quartz are characteristic of high-energy shock metamorphism and can form only under very high pressure. From experimental evidence this is not less than 30 kbar for coesite, 75 kbar for stishovite and about 100 kbar for diaplectic quartz^{9,10}. The natural occurrences confirm these high values, as coesite is found associated with astoblemes and in diamonds, and stishovite and diaplectic quartz only in astoblemes. This confirms that the single realistic hypothesis for the formation of the Vredefort Ring is by the impact of a cosmic body.

The process of formation of high-pressure silica phases is envisaged as follows. During the passage of the shock wave, a concentration of energy occurred in voids⁸ which has resulted in the formation of pseudotachylite along joints of all sizes. At the maximum pressure, diaplectic quartz formed, from which stishovite and then coesite crystallised during the progressive release of pressure. Such a process has been described for other astoblemes.

Coesite and stishovite form unusually large crystals in comparison with other known occurrences⁸. Among possible explanations, this could be attributed tentatively to a longer duration of high pressure conditions, perhaps in this case for several seconds, due to the particular magnitude of the impact. Vredefort must rank among the largest impacts known on Earth or the Moon.

The facts presented by those opposed to the impact hypothesis are pertinent, but the evidence is not strong enough to exclude the hypothesis. The metamorphic zone of contact character affecting a large part of the sedimentary strata all around the granite core^{1,11} pre-dates the formation of the shatter cones and the pseudotachylites. It has been argued that this contact metamorphism is not fortuitously associated to the Dome and indicates the presence of a large hidden batholith⁷. It would be more logical to see this metamorphism as regional, as has been suggested in the past¹¹. The presence of kyanite^{1,11}, which is an indicator of regional rather than thermal metamorphism, supports this idea. The 'contact character' particularly the development of non-shistose rocks like hornfels, could be simply due to a lack of stress, which would be expected in a thick, deeply buried sedimentary sequence without tectonic deformation (Fig. 2). This type of intracratonic metamorphism is perhaps common but its related rocks can be uplifted to the

surface in an undeformed state only by special events, such as impact phenomena.

Outside the zone affected by the astobleme, the strata of the Bushveld and Potchefstroom Basins exhibit deformation characterised by faulting, basining, arching and doming. Consequently it is probable that some components of the deformation of the Dome are not related to the impact but to normal tectonic action. For instance one effect of this action could be the skew profile of the Dome (Fig. 2c). Most of this tectonism occurred before the deposition of the little deformed Waterberg Supergroup, as a major unconformity separates the latter from the Transvaal Supergroup¹². The age of the impact (1,970 Myr) probably falls in the period of the unconformity ($\sim 2,100$ to $\sim 1,900$ Myr, ref. 13). However, there are some indications that it may have occurred after the unconformity, a hypothesis adopted in Fig. 2b, where a positive structure is postulated

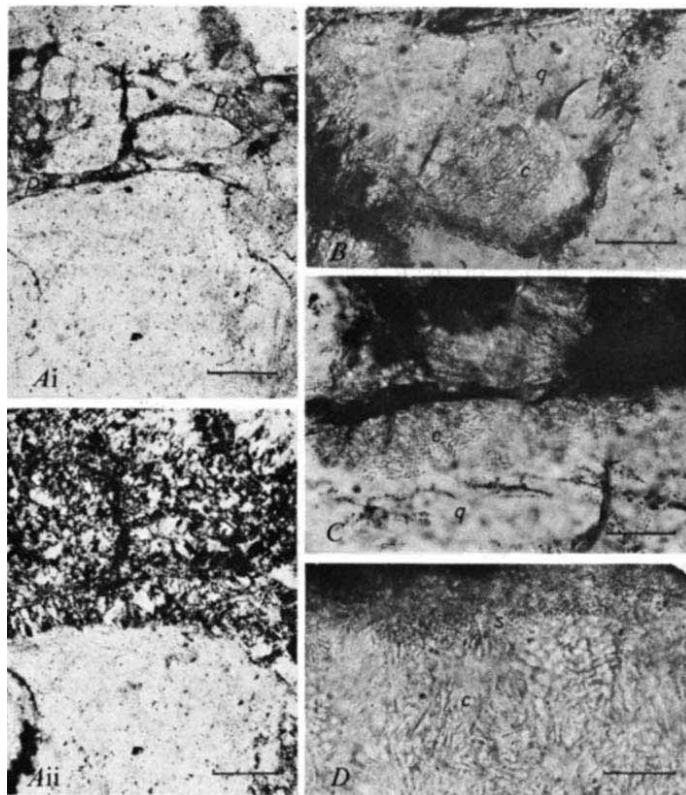


Fig. 1 A, Microbrecciated quartz in a pseudotachylite matrix (p), sample 1; A i, natural light; A ii, polarised light. Note cherty texture in the upper half representing presumed recrystallised diaplectic quartz; the lower half is untransformed monocrystalline quartz; scale bar, 200 μm . B, Radiating coesite (c), in quartz grain (q) in sample 1; scale bar, 100 μm . C, coesite (c) partly resorbed by quartz (q) in sample 1. Scale bar, 50 μm . D, Radiating and corroded stishovite growing from the edge of a former quartz grain replaced by partly reverted coesite (c); sample 1, scale bar 20 μm .

Table 2 X-ray diffraction data on stishovite

Sample 1*		ASTM 15-26	
$d(\text{\AA})$	I/I_1	$d(\text{\AA})$	I/I_1
2.95	100	2.959	100
2.25	5	2.246	18
		2.09	1
1.98	15	1.981	35
1.866	5	1.870	14
1.527	20	1.530	50
Interfere with rutile			
1.333	5	1.478	18
—		1.333	10
—		1.322	4
—		1.291	2
1.236†	5	1.235	25
1.215	5	1.215	10

*Cobalt radiation, quartz of the matrix as standard.

†Poorly visible.

30 km south-west of the impact. For example the inclination of the shatter cones in the south-east part of the Ring structure may indicate that the dip before impact was to the north-west¹⁰.

The preservation of coesite and stishovite, which is unexpected in such an old astobleme⁸, may also indicate a pre-impact tectonism. Experiments demonstrate that stishovite reverts to quartz after heating at 250 °C for 3 d (ref. 8). However, with a longer time span of geological scale, it is likely that stishovite would decompose at even lower temperature. Therefore this preservation can only have occurred if the quartzite was already at a low temperature at the time of the impact and did not undergo subsequent reheating. As the astobleme is deeply eroded, it is reasonable to assume that immediately

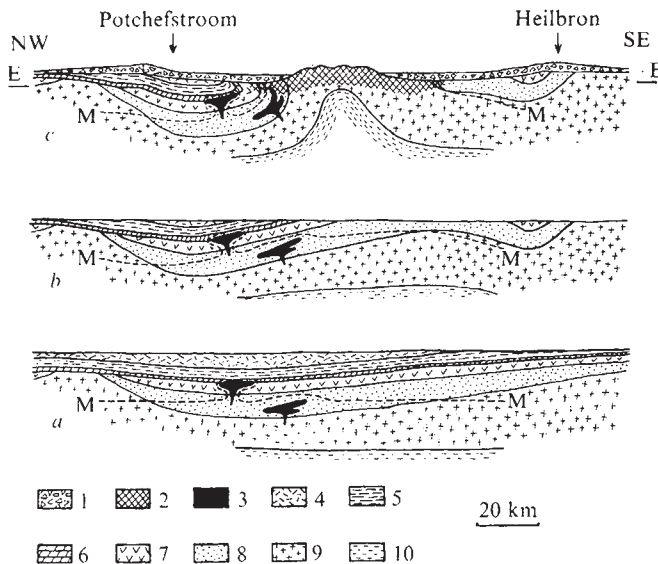


Fig. 2 Evolution of the Potchefstroom Basin and formation of the Vredefort Dome. *a*, Schematic section showing development of Potchefstroom Basin at the end of Transvaal Supergroup deposition and development of burial metamorphism, followed soon after by magmatism (only alkali granite is indicated here). *b*, Arching and doming, followed by erosion. *c*, Immediately after impact. 1, Debris due to impact; 2, strongly shattered zone; 3, alkali granite; 4, upper part of Pretoria Group (volcanics and sediments) and igneous complexes; 5, lower part of Pretoria Group; 6, Malmman Dolomite; 7, Ventersdorp Lava; 8, Witwatersrand Supergroup; 9, Archaean Granite; 10, Lower continental crust. E, approximate actual erosion surface; M, metamorphic front.

after the impact no fast cooling could have occurred in the rocks containing coesite and stishovite. This suggests that folding, uplift and erosion took place before the impact in order that the upper part of the Witwatersrand Quartzite could be raised from a depth of more than 10 km to a shallower position (Fig. 2).

We thank E. Peters kindly for assistance with X-ray diffraction analysis, B. H. Uiterwijk for photographic work and Dr R. Beeson for a critical reading of the manuscript.

J. E. J. MARTINI

Geological Survey, Private Bag X112,
Pretoria, Republic of South Africa

Received 8 December 1977; accepted 8 March 1978.

1. Nel L. T. *Geol. Surv. South Africa spec. publ.* 6, 1-130 (1927).
2. Boon, J. D. & Albritton, C. C. *Field Lab.* 5, 1-9 (1936).
3. Daly, R. A. *J. Geol.* 55, 125-145 (1947).
4. Dietz, R. S. *J. Geol.* 69, 499-516 (1961).
5. Hargraves, R. B. *J. Geol.* 78, 253-263 (1970).
6. Nicolaysen, L. O., Burger, A. J. & v. Niekerk, C. B. *Int. Un. Geodesy Geoph., 13th Gen. Ass. Berkeley* (1963).
7. Gay, N. C. *Science* 194, 724-725 (1976).
8. Stöfler, D. *J. geophys. Res.* 76, 5474-5488 (1971).
9. Kieffer, S. W. *J. geophys. Res.* 76, 5449-5473 (1971).
10. Stöfler, D. *J. geophys. Res.* 76, 5541-5551 (1971).
11. Bisschoff, A. A. thesis, Potchefstroom Univ. (1969).
12. Hall, A. L. & Molengraaff, G. A. F. *Shaler Memorial Series. Verh. d. Kon. Akad. v. Wetens. Amsterdam* 24, 1-183 (1925).
13. Walraven, F. *Abstr. 17th Congr. Geol. Soc. South Africa*, 142-145 (1977).
14. Manton, W. I. *Ann. N.Y. Acad. Sci.* 123, 1017-1049 (1965).

Cambrian trilobite diversity related to cratonic flooding

A RECENT quantitative biogeographic analysis of Cambrian trilobites¹ utilised almost all data on 4,390 genera published since 1927. The Cambrian was divided into 11 time segments (see Fig. 1) and data from over 600 papers and monographs were collated. An attempt is made here to use these data to assess the changing patterns of generic diversity

through the Cambrian. Additional data, not used in the biogeographic analysis¹ and using finer biostratigraphic subdivisions, will also be used (Fig. 3).

From Fig. 1 it is clear that the total world-wide number of genera per time segment is related to number of lists which, in turn, is related to the total number of papers published. Thus, the raw generic diversity is possibly an artefact caused by a monographic effect^{2,3}. However, the last occurrences expressed as a percentage of total genera per time segment also follow this same trend. These data have also been plotted (Fig. 2) using the method of Funnell and Cutbill⁴, who took the last occurrences in time interval one and subtracted them from the first occurrences in the next time interval. Unfortunately they did not use equal time intervals and, clearly, if a short time interval is followed by a long one, spurious extinction events might be inferred. However, if these first and last occurrences are expressed as percentages of total genera in the respective time segment, then the effects of varying time intervals will be, to some extent, overcome. This plot would then show the successive losses (extinctions) or profits relative to a base line (Fig. 2). The method suggests that the major extinction event occurred at the end of the early Cambrian, with minor extinctions between time segments 8 and 9 and at the end of the Trempealeau.

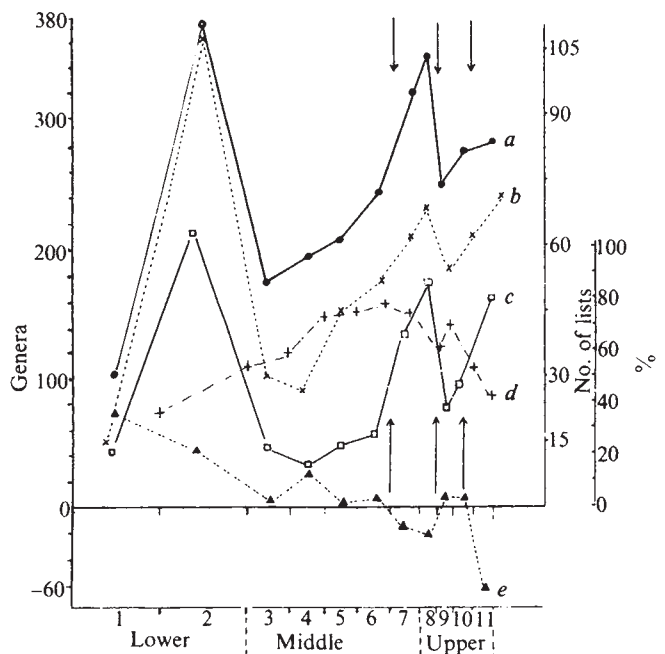


Fig. 1 Plot of total diversity through the Cambrian (*a*). Last occurrences (*c*) are expressed as a percentage of total genera present per time segment. The number of faunal lists (*b*) is closely related to the number of papers written. Arrows indicate position of biore boundaries. *d*, No. of survivors from previous time segment. *e*, Net per time segment. Time segments are as follows: 1, Lower Olenellus Zone; 2, Upper Olenellus Zone; 3, Plagiura Zone; 4, Albertella Zone; 5, Glossopleura Zone; 6, Bathyriscus-Elrathina Zone; 7, Bolaspidea Zone; 8, Cedaria & Crepicephalus Zones; 9, Aphelaspis & Dunderbergia Zones; 10, Elvinia & Conaspis Zones; 11, Ptychaspis, Prosaikia & Saukia Zones. Correlations outside North America based on correlation chart of Cowie *et al.*¹⁵

McAlester⁵ used a plot of first occurrences minus last occurrences expressed as a percentage of total genera present per time segment. This method (Fig. 2) reveals a general decline through the Cambrian.

When more refined data are used for one major tectonic block (North America) a good correlation is found between the extent of cratonic flooding and last occurrences of

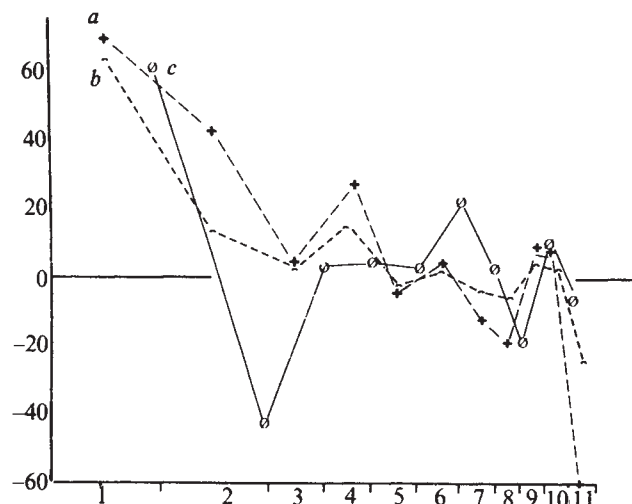
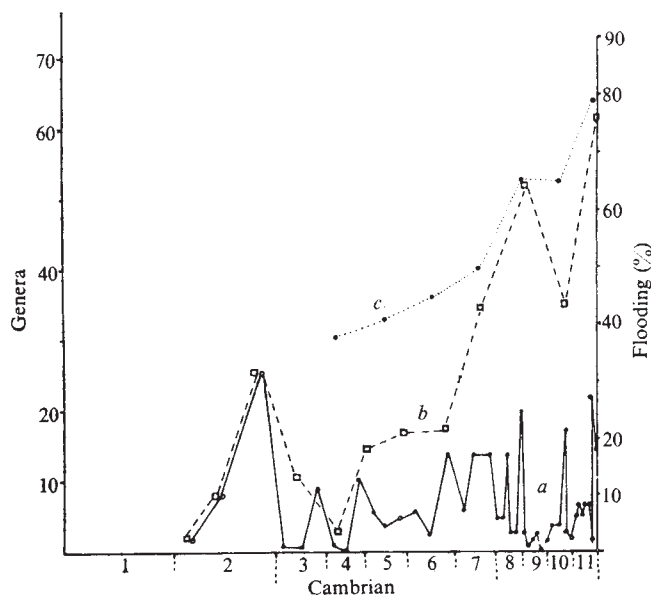


Fig. 2 Plots of total world-wide Cambrian trilobite genera. *a*, Net: first occurrences of genera minus last occurrences per time segment. *b*, Net as % total: firsts minus lasts as a percentage of total genera per time segment. *c*, Profit and loss: last occurrences in time interval one subtracted from first occurrences in next time interval, both expressed as % of total genera present.

genera (Figs 3 and 4a). There is also a very strong correlation between total diversity of endemic North American genera and cratonic flooding (Fig. 4b). Ideally, we should calculate the extent of cratonic flooding for every tectonic block, but the data for the USSR⁶ and Australia⁷, although reliable, are based on very coarse time units. No reliable palaeogeographic data are available for other parts of the world or, if they exist do not attempt any subdivision of the Cambrian. However, it is interesting that both the total world diversity of cosmopolitan genera (Fig. 4c) and provincial genera (Fig. 4d) show strong correlations with North American cratonic flooding. If these relationships are taken to exhibit an area effect, that is, the diversity is related to area⁸⁻¹¹, then the correlation of diversity to area is easily explained. However, the correlation of last occurrences to extent of cratonic flooding (Fig. 4a) is not as easily explained nor is it theoretically expected. As flooding in-

Fig. 3 Plot of last occurrence of North America (tectonic block) genera. These are not necessarily endemic genera. *a*, A plot of the raw data which were compiled, using the finest possible biostratigraphic subdivisions, from a large number of papers published from 1950-76. *b*, The same set of data plotted as equal time units. *c*, Plot of North American cratonic flooding using very similar time units to *b*. Flooding data based mainly on Lochman¹³.



creases, greater areas are opened up for colonisation. Niche partitioning would presumably take place and, therefore, the total diversity would increase. With a greater number of taxa, competition would increase and, therefore, last occurrences would also increase. During a regression, such as that during the Permian, last occurrences of shelf benthos would increase but the level of first occurrences would be very low and hence there would be a gradual but catastrophic net extinction event¹¹. The net extinction events shown in Fig. 2 may be due to rapid, though minor, regressions. An early Medial Cambrian regression is suspected but is not well documented. The extinctions between time segments 8 and 9 (Fig. 2) are related to a well documented regression¹³ (Fig. 3), whilst the net extinctions in

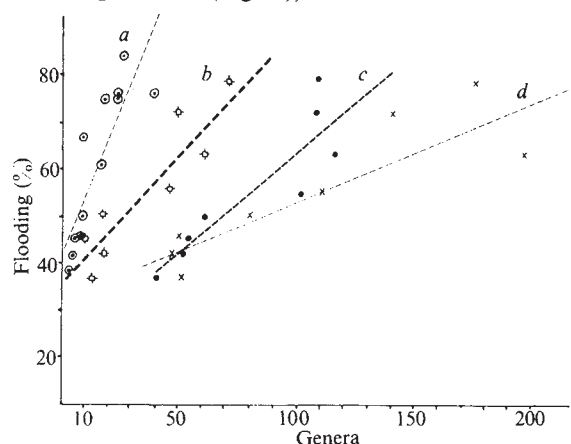


Fig. 4 Number of genera plotted against North American cratonic flooding. *a*, Last occurrences found on North American tectonic block ($r = 0.837$). *b*, North American block endemic diversity. ($r = 0.905$). *c*, Total world cosmopolitan diversity ($r = 0.87$). *d*, Total world provincial diversity ($r = 0.87$).

the Trempealeau could be due to inadequate data on generic ranges across the ill-defined and poorly correlated Cambro-Ordovician boundary. Extinctions across biomere boundaries¹⁴ are apparent in Fig. 1 (worldwide) and Fig. 3 (North America) but, except for the extinction event between time segments 8 and 9, are not particularly significant on the time scale used here.

As cratonic flooding increased during the Cambrian, so would the possibility of preserving fossiliferous outcrops leading to the observed increase in the number of published papers (Fig. 1). A real increase in diversity through the Cambrian would also result in an increase in papers and monographs.

Taylor¹² has calculated the average duration of biostratigraphic subdivisions through the Cambrian as follows: early, 7.5×10^6 yr; medial 7.5×10^6 , and late 0.94×10^6 . Slightly different durations have been calculated for Fig. 1. This increase is very probably a result of the increase in last occurrences during the late Cambrian.

Thus, as theoretically expected, generic diversity of trilobites is positively correlated to cratonic flooding. More surprisingly, however, it shows that the number of published papers, the number of last occurrences and the precision of biostratigraphic subdivisions are also positively correlated to cratonic flooding.

The Australian Research Grants Committee supplied financial support and A. Duncan and B. Stait made helpful comments.

CLIVE F. BURRETT
ROBERT G. RICHARDSON

Geology Department,
University of Tasmania,
Box 252C Hobart,
Tasmania, Australia 7001

Received 6 January; accepted 30 January 1978.

1. Burrett, C. & Richardson, R. *Tectonophysics* (submitted).
2. Raup, D. M. *Paleobiology* 3, 328-329 (1977).

3. Sheehan, P. M. *Paleobiology* 3, 325–327 (1977).
4. Funnell, B., & Cutbill, C. in *The Fossil Record* (ed. Harland, W. B.) 623 (Geological Society, London, 1967).
5. McAlester, A. L. in *The Permian and Triassic Systems and Their Mutual Boundary* 11–15 (Canadian Society of Petroleum Geologists, Alberta, 1973).
6. Vinogradov, A. P. (ed.) *Atlas of the Palaeogeographic-Lithologic Maps of the U.S.S.R.* (Academy of Science U.S.S.R., Moscow, 1968).
7. Webby, B. D. *Univ. Sydney Repts Geol. Sci.* 3, 1–95.
8. Schopf, T. J. M. *J. Geology* 82, 267–274.
9. Simberloff, D. S. in *Models in Paleobiology* (ed. Schopf, T. J. M.) 267–286 (Wiley, New York, 1972).
10. MacArthur, R. H. *Geographical Ecology* (Harper & Row, New York, 1972).
11. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton University Press, 1967).
12. Taylor, M. in *Concepts and Methods in Biostratigraphy* (eds Kauffmann, E. G. & Hazel, J. E.) 397–425 (Dowden, Hutchinson & Ross, Stroudsburg, 1977).
13. Lochman, C. *Geol. Soc. Am. Bull.* 81, 149–153 (1970).
14. Palmer, A. R. *J. Paleontol.* 39, 149–153 (1965).
15. Cowie, J. W., Rushton, A. W. A. & Stubblefield, C. J. *spec. Rep. Geol. Soc. London* 2, 1–43 (1972).

Larvicidal activity of diflubenzuron in the tsetse fly

THE insecticides used in control operations against tsetse flies (*Glossina* spp.) have so far been almost exclusively organochlorine compounds, which kill the adult insect. Because of the possibility of the development of resistance to these compounds and as they may eventually become unavailable or unacceptable because of their persistence in the environment, a variety of possible new insecticides for tsetse control have been evaluated¹. Dimilin, diflubenzuron, (*N*-{(4-chlorophenyl)amino}carbonyl}-2,6-difluorobenzamide), is one of a range of insect growth regulators (IGRs) which interfere with chitin synthesis^{2,3} and, specifically, when applied to the adult female, inhibit egg hatch. Such effects have been reported for many pest insects, including a number of oviparous species of Diptera such as the house fly, *Musca domestica* L.⁴, stable fly, *Stomoxys calcitrans* L.⁴, horn fly, *Haematobia irritans* L.⁵, black fly, *Simulium vittatum* Zetterstedt⁶, various mosquitoes⁷ and chironomid midges⁸. Results reported here indicate the potential of a new range of compounds such as diflubenzuron which might be employed against *Glossina* in the future and which, following application to the adult female fly, act as a larvicide.

Glossina spp. are larviparous, and have a long reproductive life, a single egg being ovulated at approximately 9-d intervals at 25 °C, and embryogenesis and larval development taking place in the uterus. After larviposition, the third instar larva does not feed and, in nature, rapidly burrows into the soil where it forms a puparium. Experimental flies were obtained from the colony of *G. morsitans morsitans* Westw. at this laboratory, where the maintenance conditions are 25 ± 0.5 °C and 60–70% relative humidity; flies were given the opportunity of feeding on the flank of a goat on 6 d a week⁹.

Four groups of 25 female *G. m. morsitans* were mated when 3 d of age and, on day 4, treated topically on the thorax with 1-μl solutions of diflubenzuron in acetone at rates of 0.05, 0.5, 1.0 and 5.0 μg per fly. The flies were subsequently held in cages, and mortality and numbers of offspring were recorded over 12 (9-d) age groups; optimal fecundity would have been one offspring per female per age group. At 5 μg per fly there was no effect on either female longevity or the numbers of offspring produced, compared to controls (Fig. 1). However, whereas 97% of the 159 control offspring pupariated normally, 96% of the 173 offspring of treated females failed to pupariate; 17 of these were aborted as first or second instar larvae, but all the others were apparently full-term third instar larvae which had been produced normally and then had failed to pupariate. At lower dose rates of diflubenzuron (Fig. 1) a greater number of apparently normal offspring were produced, especially towards the end of the reproductive life of the females, but even 0.05 μg per fly affected offspring produced in age group 2, the first productive age group. Some of the apparently normal puparia of treated females were affected, as they failed to produce an adult fly; emergence rates from total offspring of females treated with the various doses of diflubenzuron (Table 1)

were significantly different from one another (by χ^2 ; $P < 0.001$ in all cases), except rates from puparia of females treated with 1.0 μg or 0.5 μg.

Two groups of 25 females treated topically with diflubenzuron on day 18, after the production of one normal offspring, subsequently produced abnormal offspring in equivalent numbers to those produced by 3-d-old females treated with the same doses of diflubenzuron. Thus, females treated with 5 μg per fly produced, from age group 3 onwards, one apparently normal puparium and 111 abnormal offspring; females treated with 1 μg per fly produced 25 apparently normal puparia and 114 abnormal offspring.

Four groups of 25 female *G. m. morsitans* were mated when 3-d-old and then, on day 4, brought into tarsal contact with a surface treated with diflubenzuron. One group was exposed for 30 s to a glass fibre disk (20 mm diameter) coated with a suspension of diflubenzuron in acetone such that after solvent evaporation there was 1.6 mg compound cm⁻²; 71 apparently normal puparia and 39 abnormal offspring were produced in age groups 2–6 (1 normal : 20 abnormal in age group 2); thereafter, all offspring were unaffected. The other three groups of females were exposed in a WHO insecticide test kit¹⁰ to Whatman no. 1 filter paper soaked with a solution of diflubenzuron in

Fig. 1 Longevity (graph) and production of apparently normal puparia (shaded histogram) and abnormal offspring (open histogram) by female *G. m. morsitans* following topical application of various doses of diflubenzuron. a, 5.0 μg per fly; b, 1.0 μg per fly; c, 0.5 μg per fly; d, 0.05 μg per fly; e, control.

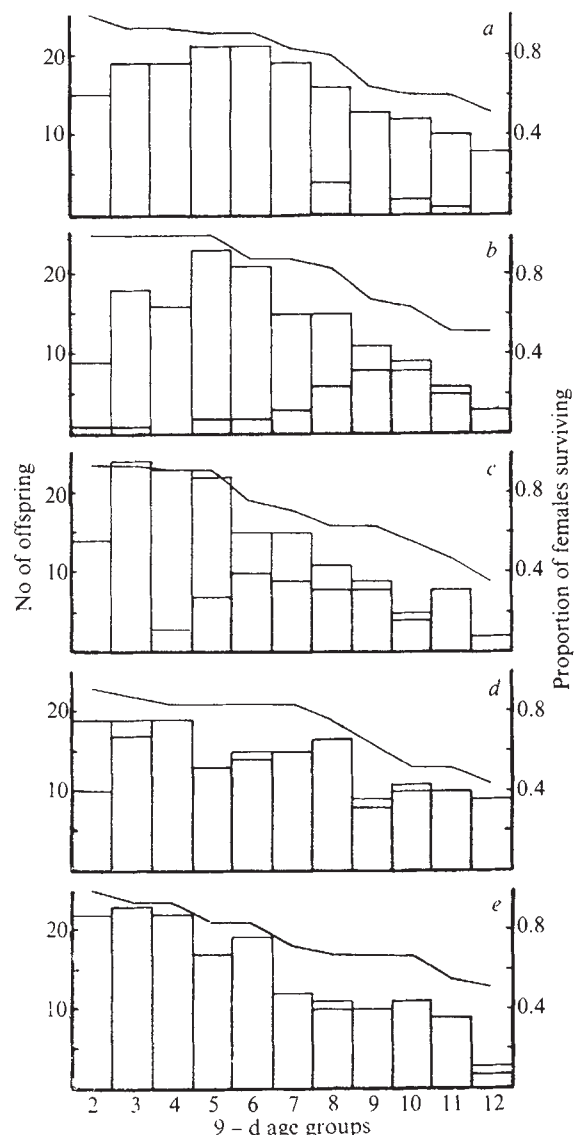


Table 1 Adult emergence rate from total offspring of female *G. m. morsitans* after topical application of various doses of diflubenzuron

Dose (μg fly)	Total offspring	No. adults emerged
5.0	173	5 (3%)
1.0	146	29 (20%)
0.5	148	25 (17%)
0.05	156	132 (85%)
Control	159	154 (97%)

N, *N*-dimethylformamide, acetone and olive oil such that after solvent evaporation there was 1.1 mg compound cm^{-2} on oily paper. Following exposure for 2, 1 or a $\frac{1}{2}$ h the ratio of apparently normal to affected offspring, over age groups 2–4, was 24 : 25, 39 : 20 and 23 : 22, respectively; thereafter, only normal offspring were produced. Again there was a decline in the numbers of abnormal offspring after age group 2, at which time virtually all offspring were affected.

Twenty-five 7-d-old male *G. m. morsitans* treated with 5 μg per fly of diflubenzuron transferred some of the compound to females during mating 24 h later, but only a few affected larvae were produced and only during the first productive age group (16 apparently normal : 6 affected). The same males produced no effect after mating with new females one week later.

Clearly, in our experimental conditions, the most effective route of diflubenzuron into female *G. m. morsitans* was by topical application, although less pronounced effects followed tarsal contact with deposits and, to a much lesser extent, contact during mating with treated males.

In oviparous insects, diflubenzuron applied to the adult female inhibits egg hatch, but in the larviparous tsetse fly, eggs hatch normally and development is not interfered with until the time of pupariation. It can be speculated that no chitin synthesis is necessary in the uterine environment of the early developmental stages and that such synthesis is only required to provide protection for the first free-living stage.

Topical application of 10 μg of a juvenile hormone analogue induces abortion in *Glossina*, but virtually only in the current reproductive cycle¹¹. However, a single application of diflubenzuron of about 0.5 μg per fly results in the production of mostly nonviable offspring throughout a female reproductive life of more than 100 d. Although some normal puparia are produced towards the end of reproductive life, they would certainly alone be insufficient to maintain a natural population. Diflubenzuron thus shows promise as a potential agent for the control of *Glossina* in the field. It has negligible vertebrate toxicity and can be appropriately formulated for aerial application of fine droplets, when it can have equivalent effects to those of similar quantities of DDT¹².

We thank Dr A. B. Borkovec, Agricultural Research Center (USDA), Beltsville for proposing this investigation and for supplies of diflubenzuron, Dr P. A. Langley for helpful comments and the Ministry of Overseas Development for financial support.

A. M. JORDAN

M. A. TREWERN

*Tsetse Research Laboratory, University of Bristol,
School of Veterinary Science, Langford, Bristol, UK*

Received 1 December 1977; accepted 6 March 1978.

- Hadaway, A. B., Barlow, F., Turner, C. R. & Flower, L. S. *Pestic. Sci.* **8**, 172–176 (1977).
- Wellington, K., Mulder, R. & Daalen, J. J. van *J. agric. Fd Chem.* **21**, 348–354 (1973).
- Post, L. C. & Vincent, W. R. *Naturwissenschaften* **60**, 431–432 (1973).
- Wright, J. E. & Spates, G. E. *J. econ. Ent.* **69**, 365–368 (1976).
- Wright, J. E. & Harris, R. L. *J. econ. Ent.* **69**, 728–730 (1976).

- Lacey, L. A. & Mulla, M. S. *J. econ. Ent.* **70**, 369–373 (1977).
- Mulla, M. S. & Darwazeh, H. A. *J. econ. Ent.* **69**, 309–312 (1976).
- Mulla, M. S., Kramer, W. L. & Barnard, D. R. *J. econ. Ent.* **69**, 285–291 (1976).
- Nash, T. A. M., Jordan, A. M. & Boyle, J. A. *Ann. trop. Med. Parasit.* **62**, 336–341 (1968).
- Tech. Rep. WHO* No. 191, (1960).
- Denlinger, D. L. *Nature* **253**, 347–348 (1975).
- Neisess, J., Markin, G. P. & Schaefer, R. J. *J. econ. Ent.* **69**, 783–786 (1976).

Complementation analysis of hormone-sensitive adenylate cyclase

PLASMA membranes of most animal cells respond to specific hormonal signals by synthesising adenosine 3',5' cyclic monophosphate (cyclic AMP) from ATP, a reaction catalysed by adenylate cyclase. The precise molecular basis of this membrane response has not been defined. In particular, we do not know the number of molecular species required for transduction of hormonal signals into stimulation of cyclic AMP synthesis. Catalytic adenylate cyclase and specific hormone binding activities can be separated on the basis of their physical properties^{1–3}. However, hormonal stimulation may require one or more other molecules necessary for coupling the two activities in membranes. One of these may contain a guanyl nucleotide regulatory site, as guanyl trinucleotides are strictly required for hormonal stimulation of cyclase in several membrane systems^{4–8}. Genetic complementation analysis can help to define the minimum number of gene products involved in a complex biochemical function. Although subject to qualifications, the principle is simple: if a cross between two mutant cells bearing a recessive phenotype fails to reconstitute the wild-type (WT or normal) phenotype, both mutant parental cells must bear a lesion in a common gene product; if the cross produces a WT phenotype, the two parental mutations must affect different gene products. Here we apply this method of analysis to the hormone-sensitive cyclase system of the S49 lymphoma tissue culture line, in which two distinct classes of variant clones with stable deficiencies in hormonal stimulation of cyclase have been selected^{9–10}. We have constructed viable hybrid clones by crossing each of these variants with WT, and with each other. We describe the resulting hybrid phenotypes and their implications regarding the molecular components of hormone-sensitive adenylate cyclase.

Table 1 summarises the genetic properties, nomenclature, and mode of selection of the S49 clones used in these experiments. WT S49 cells accumulate cyclic AMP in response to β -adrenergic agents, prostaglandin E_1 (PGE_1), and cholera enterotoxin⁹. In addition to these agents, NaF and guanylyl-5'-imidodiphosphate (Gpp(NH)p) stimulate adenylate cyclase in WT membrane preparations^{9,11}. In contrast, none of these five classes of effectors increase cyclic AMP synthesis in cells or membranes of the cyc^- phenotype (previously termed AC^- , refs 9 and 12). A second type of S49 variant, recently reported by Haga *et al.*¹⁰, exhibits defective coupling between hormone receptors and cyclase: in comparison with WT, cyclase in the 'uncoupled' (UNC) variant shows a reduced response to PGE_1 , and weakly responds to isoprenaline, a β -adrenoceptor-agonist, in the presence of Gpp(NH)p but not GTP. The UNC cyclase, however, can be stimulated by cholera toxin, guanyl nucleotides, and NaF. Both cyc^- and UNC cells bear β -adrenoreceptors, detectable by binding of a radio-labelled β -adrenoreceptor antagonist, that are similar in number and stereoselectivity to the β -adrenoreceptors of WT^{10,12}. We tested parental and hybrid cells (Fig. 1) or membrane preparations (Fig. 2) for stimulation of adenylate cyclase by various effectors.

In crosses with WT, the loss of hormonal stimulation of cyclic AMP synthesis in both cyc^- and UNC cells behaved as a recessive character: $\text{WT} \times \text{cyc}^-$ and $\text{WT} \times \text{UNC}$ hybrid

cells and/or membranes responded like WT to all five effectors that stimulate cyclic AMP synthesis, including isoprenaline and PGE₁.

Qualitatively, these results were highly reproducible. At least two independently selected hybrid clones of each type were tested, and each result in Figs 1 and 2 was repeated three to six times, with identical patterns in each case. Quantitatively, the hormonal responses of these hybrids seemed to be generally intermediate between those of the WT and variant parents (Figs 1 and 2), although (as previously reported^{8,16}) there is considerable interexperimental variation in the magnitude of cyclic AMP responses (assayed in whole cells or membrane preparations), even in WT. WT × WT hybrids showed a WT pattern of response, indicating that the hybridisation procedure *per se* did not alter the regulation of cyclase. *cyc*⁻ × *cyc*⁻ hybrid cells showed a *cyc*⁻ phenotype (Fig. 1); the failure of these two independently selected *cyc*⁻ clones to complement each other indicates that they bear a similar or identical lesion.

The recessive character of both *cyc*⁻ and UNC phenotypes in crosses with WT is consistent with the view that each variant has lost the function of some component or components of the hormone-sensitive cyclase system. This result indicates that neither of the variant phenotypes is produced by the presence of some diffusible inhibitor of cyclase or of adrenoreceptor-cyclase coupling. The two variant phenotypes are similar only with regard to loss of hormonal stimulation of cyclase; catalytic activity of UNC cyclase, unlike *cyc*⁻, is stimulated by cholera toxin, guanyl nucleotide and NaF. These differences suggested that the UNC and *cyc*⁻ phenotypes would complement each other in a hybrid containing both genomes; as each lesion acted as a recessive in crosses with WT, each of the variants should be able to supply a component lacking in the other, and the UNC × *cyc*⁻ hybrid should exhibit a WT response to hormone.

Surprisingly, however, UNC × *cyc*⁻ hybrid cells and membranes failed to synthesise cyclic AMP in response to isoprenaline (Figs 1 and 2). The UNC × *cyc*⁻ hybrid pheno-

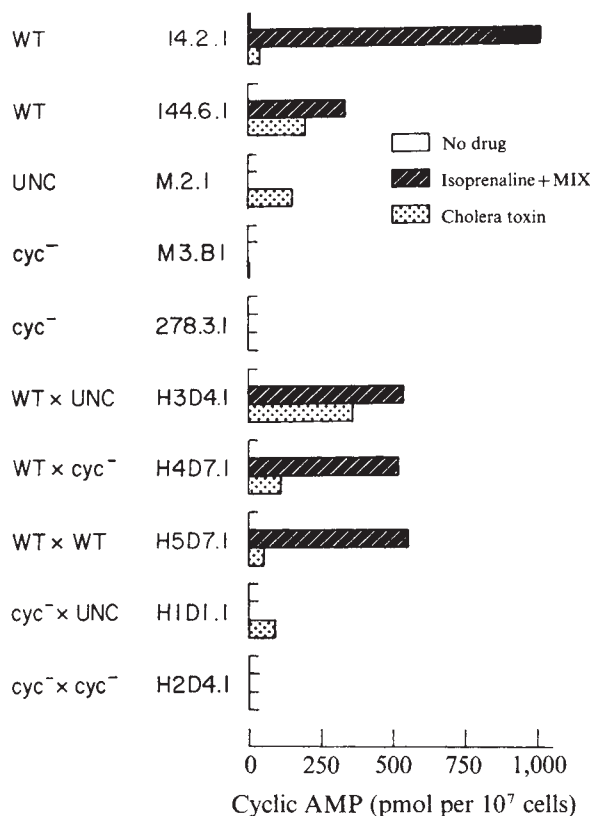


Fig. 1 Cyclic AMP accumulation in cells. Cells were exposed to isoprenaline 10 μ M + 1-methyl-3-isobutyl-xanthine (MIX) 500 μ M for 10 min or to cholera toxin 100 ng ml⁻¹ + MIX 500 μ M for 90 min. The cells were centrifuged and assayed for cyclic AMP by the competitive protein binding method of Gilman¹⁵.

type was similar to that of the UNC parent, both in its response to cholera toxin, guanyl nucleotide and NaF, and in showing a small response to PGE₁ (Fig. 2).

The unexpected failure of the UNC and *cyc*⁻ lesions to complement each other strongly suggests that both these variants bear a common functional deficiency in the mechanism that couples hormone receptors and adenylate cyclase in WT cells. This conclusion is further supported by parallel *in vitro* reconstitution experiments using the same cell lines, recently reported by Ross and Gilman¹⁷. These workers found that addition of detergent-solubilised extracts of WT membranes to *cyc*⁻ membranes resulted in reconstitution of adenylate cyclase activity that was responsive to hormones as well as guanyl nucleotides and NaF; similarly, solubilised WT extracts conferred on UNC membranes the ability to synthesise cyclic AMP in response to hormones. However, solubilised UNC membrane extracts did not reconstitute the hormonal response when added to *cyc*⁻ membranes, although guanyl nucleotide and NaF responses were reconstituted. We have confirmed these results (G. L. J. *et al.*, unpublished data). Thus, both in membranes and in viable hybrid cells, the *cyc*⁻ and UNC phenotypes behave as recessives with respect to WT, but the two variant lesions fail to complement each other.

Taken alone, without the hybrid complementation and *in vitro* reconstitution results, the phenotypes of WT and variant S49 clones suggest that hormone-sensitive adenylate cyclase must be made up of at least three components: hormone receptors, present in WT and both variant cell types; a 'coupling' molecule, defective in UNC (but perhaps not completely absent, as responses to both PGE₁ and β -adrenergic amines are detectable¹⁰); and catalytic adenylate cyclase itself. We initially proposed⁹ that the lesion responsible for the *cyc*⁻ phenotype was a loss of catalytic adenylate cyclase. Several observations, however,

Table 1 Parent cell genetic markers and hybrids formed

Parents		Cross	Hybrids
Phenotype	Clone number		
WT, TK ⁻	14.2.1	WT × UNC (14.2.1 × M.2.1)	H3D4.1
WT, HPRT ⁻	144.6.1	WT × <i>cyc</i> ⁻ (144.6.1 × M3.B1)	H4D7.1
UNC, HPRT ⁻	M2.1	WT × WT (14.2.1 × 144.6.1)	H5D7.1
<i>cyc</i> ⁻ , TK ⁻	M3.B1	<i>cyc</i> ⁻ × UNC (M3.B1 × M2.1)	H1D1.1
<i>cyc</i> ⁻ , HPRT ⁻	278.3.1	<i>cyc</i> ⁻ × <i>cyc</i> ⁻ (M3.B1 × 278.3.1)	H2D4.1

All lines were near-isogenic subclones of the S49 lymphoma tissue culture cell. WT and thymidine kinase-deficient (TK⁻) cells were obtained from the Salk Institute and UNC from A. G. Gilman. Lines with multiple genetic markers were produced by sequential single-step cloning in agar containing the indicated selective agents. TK⁻ cells were resistant to bromouridine deoxyribose 30 μ M; hypoxanthine phosphoribosyltransferase-deficient (HPRT⁻) cells were resistant to 6-thioguanine 10 μ M; *cyc*⁻ cells were selected in cholera toxin 100 ng ml⁻¹ + Ro20-1724 (4-(3-butoxy-4-methoxybenzyl)-2-imidazolidinone) 30 μ M, which yielded only *cyc*⁻ mutants, or in TRM medium (terbutaline 100 μ M + Ro20-1724 30 μ M + 1-methyl-3-isobutylxanthine (MIX) 100 μ M)¹⁰, which yielded UNC as well as *cyc*⁻ mutants. Hybrids were formed essentially as described by Geffer *et al.*¹³ using polyethylene glycol (PEG) 1000 as the fusing agent and hypoxanthine-aminopterin-thymidine (HAT) medium as the selective agent¹⁴. 5 × 10⁶ cells of each parent line were used per fusion. The TK⁻ and HPRT⁻ cells were non-revertible (<10⁻⁸ survival in HAT medium). The hybrid nature of the cells was established by their ability to form colonies in HAT and by their DNA content (determined by flow microfluorimetry¹⁴) and chromosome number which were both twice that of the pseudodiploid parents. No diminution in DNA or chromosome content of the hybrids was detected in the period required to do the experiments.

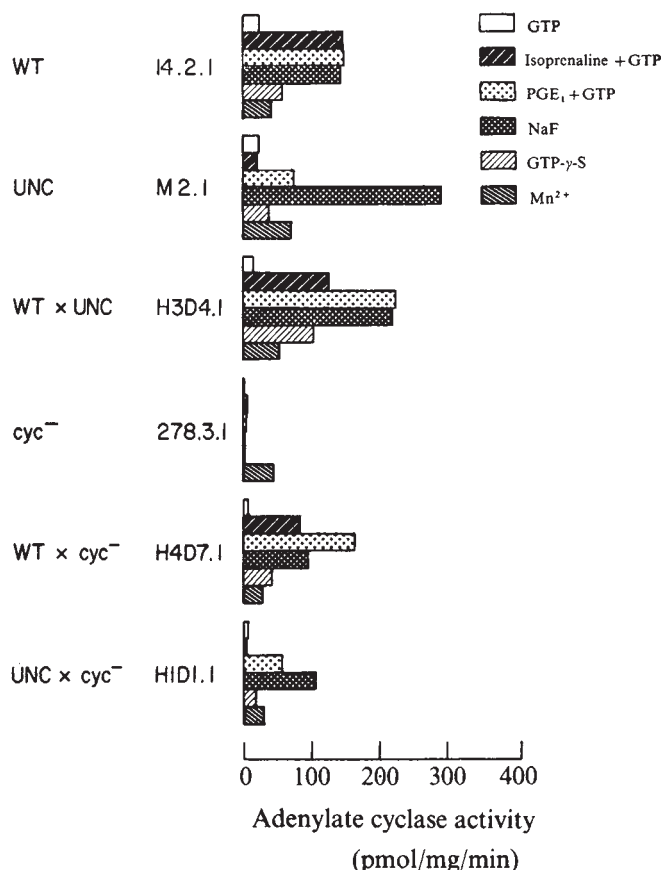


Fig. 2 Adenylate cyclase activity of membrane preparations. Plasma membrane purification was accomplished and adenylate cyclase activity was measured by the method of Ross *et al.*⁸ with modifications¹¹. Enzyme activity was measured in a volume of 100 μ l with ATP 1 mM, Na HEPES 50 mM, 1-methyl-3-isobutyl-xanthine (MIX) 0.2 mM, 2-mercaptoethanol 5 mM and bovine serum albumin 0.2 mg ml⁻¹ at pH 8.0. The buffer solution also contained MgCl₂ 4 mM unless this salt was replaced by the same concentration of MnCl₂ (as indicated). Other effectors included GTP 0.1 mM, isoprenaline 10 μ M + GTP 0.1 mM, PGE₁ 100 μ M + GTP 0.1 mM, NaF 10 mM, or guanosine-5'-(γ -thio)triphosphate (GTP γ S) 0.1 mM. GTP γ S is a GTP analogue relatively resistant to hydrolysis, which produces quasi-irreversible stimulation of adenylate cyclase in avian erythrocytes¹⁹ and S49 membranes (unpublished) qualitatively similar to the effects of Gpp(NH)p. Reactions were initiated by the addition of approximately 10 μ g of membranes and continued for 10 min at 30 °C, with shaking, and stopped by addition of 50 μ l 150 mM acetic acid. Cyclic AMP was measured as in Fig. 1.

suggest that this is not the case: cyc⁻ cells do contain cyclic AMP and cyc⁻ membranes synthesise cyclic AMP, although both are low in comparison with WT values⁹; synthesis of cyclic AMP can be demonstrated with both WT and cyc⁻ membranes when Mn²⁺ is substituted for Mg²⁺ in the reaction buffer (Fig. 2; E. Ross, personal communication); in the *in vitro* reconstitution of adenylate cyclase activity, using detergent-solubilised extracts from cyc⁻ and WT membranes, cyc⁻ can be shown to contribute a component that is similar in its thermal stability and inactivation by *N*-ethylmaleimide to WT catalytic activity¹⁸. Further experiments will be necessary to determine whether the catalytic component stimulated by hormones in WT is identical to the Mn²⁺-sensitive cyclase present in both WT and variant membranes. If so, we must postulate a fourth component of hormone-sensitive adenylate cyclase, the alteration of which is responsible for the cyc⁻ phenotype.

The introduction of a fourth component can be reconciled with the failure of cyc⁻ and UNC to complement one another if the cyc⁻ phenotype functionally lacks at least two gene products, one of which is also deficient in UNC. This

interpretation would require coordinate loss of two membrane components in cyc⁻ produced by some mechanism that does not interfere with expression of WT components in WT x cyc⁻ hybrids, as these hybrids show the WT phenotype (Figs 1 and 2).

We propose an alternative, simpler interpretation of the lack of complementation. Hormone-sensitive adenylate cyclase is considered to be composed of a minimum of three components—receptors, catalytic cyclase, and a single class of coupling molecules—that are probably involved in regulation of adenylate cyclase by guanyl nucleotides. The coupling molecule is necessary for stimulation of catalytic cyclase by all effectors except Mn²⁺, and is functionally absent in cyc⁻. In this view, the UNC phenotype could be produced by an alteration of that same coupling molecule or by a reduction in its concentration in membranes. In either case, such a change in the coupling molecule could impair its interaction with hormone receptors, but allow it to mediate cyclase activation by guanyl nucleotide, cholera toxin and NaF.

These interpretations can be tested, using UNC and cyc⁻ membranes as assay systems for the separation, purification, and biochemical characterisation of the WT membrane component or components required for reconstituting sensitivity of the variant membranes to hormones.

We thank Mary Gleason, Michele Sanda and Theodoore van Daalen Wetters for technical assistance. This work was supported by the USPHS and the NSF. P.C. is the recipient of a Research Career Development Award from the USPHS.

JOSEFINA NAYA-VIGNE*†

GARY L. JOHNSON*

HENRY R. BOURNE*†§

PHILIP COFFINO*†

Division of Clinical Pharmacology, Department of Medicine*,
Department of Microbiology†,
Cardiovascular Research Institute‡, and
Department of Pharmacology§,
University of California,
San Francisco, California 94143

Received 27 October, 1977; accepted 13 February, 1978.

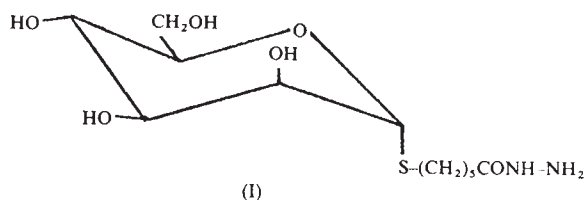
1. Limbird, L. E. & Lefkowitz, R. J. *J. biol. Chem.* **252**, 799–802 (1977).
2. Haga, T., Haga, K. & Gilman, A. G. *J. biol. Chem.* **252**, 5776–5782 (1977).
3. Welton, A. F. *et al. J. biol. Chem.* **252**, 5447–5450 (1977).
4. Rodbell, M., Birnbaumer, L., Pohl, S. L. & Krans, H. M. J. *J. biol. Chem.* **246**, 1877–1882 (1971).
5. Rodbell, M., Lin, M. C. & Salomon, Y. *J. biol. Chem.* **249**, 59–65 (1974).
6. Bilizekian, J. P. & Aurbach, G. D. *J. biol. Chem.* **249**, 157–161 (1974).
7. Londos, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 3087–3090 (1974).
8. Ross, E. M., Maguire, M. E., Sturgill, T. W., Biltonen, R. L. & Gilman, A. G. *J. biol. Chem.* **252**, 5761–5775 (1977).
9. Bourne, H. R., Coffino, P. & Tomkins, G. M. *Science* **187**, 750–752 (1975).
10. Haga, T. Y., Ross, E. M., Anderson, H. J. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2016–2020 (1977).
11. Johnson, G. L. & Bourne, H. R. *Biochem. biophys. Res. Commun.* **78**, 792–798 (1977).
12. Insel, P. A. *et al. Molec. Pharmacol.* **12**, 1062–1069 (1976).
13. Geffer, M. L., Margulies, D. H. & Scharff, M. D. *Somatic Cell Genet.* **3**, 231–236 (1977).
14. Lemaire, I. & Coffino, P. *J. Cell Physiol.* **92**, 437–446 (1977).
15. Gilman, A. G. *Proc. natn. Acad. Sci., U.S.A.* **67**, 305–312 (1970).
16. Bourne, H. R., Coffino, P. & Tomkins, G. M. *J. Cell Physiol.* **85**, 611–620 (1975).
17. Ross, E. M. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3715–3719 (1977).
18. Ross, E. M. & Gilman, A. G. *J. biol. Chem.* **252**, 6966–6969 (1977).
19. Cassel, D. & Selinger, Z. *J. cyclic Nucleotide Res.* **3**, 11–22 (1977).

Synthetic concanavalin A receptors and erythrocyte agglutination

THERE is considerable evidence to suggest the involvement of cell surface carbohydrates in cell-cell recognition and adhesion in, for example, muscle cell fusion¹ and the adherence of bacteria to cells². It is not yet clear whether the selectivity of the recognition process is a consequence of the structure and sequence of the saccharide chain and/or the topographical distribution of the saccharide unit(s) in the membrane. Further insight into the relationship between cellular specificity and carbohydrate structure

and distribution may be gained if it were possible to introduce saccharide molecules on to specific portions of the cell surface. We show here that bovine erythrocytes, which are not normally agglutinated by the plant lectin concanavalin A (con A) (refs 3, 4), can be rendered agglutinable by the chemical introduction on to the cell surface of α -D-mannose residues, which are thought to be the primary specificity determinant for this lectin. The lectin-induced agglutination of cells can be taken as a model for the investigation of the general problem of the role of sugars in cell-cell recognition.

We have prepared a hydrazide derivative (I) of a thioglycoside of α -D-mannose by the reductive alkylation of 2-S-(2,3,4,6-tetra-*O*-acetyl- α -D-mannopyranosyl)-2-thiopseudourea hydrobromide⁵ with bromohexanoic acid, followed by carbodiimide-mediated coupling to *t*-butylcarbazate, subsequent deacetylation with sodium methoxide



and removal of the *t*-butoxycarbonyl group with anhydrous HCl/methanol. This hydrazide can be covalently attached to aldehydic groups generated on the intact erythrocyte membrane by mild periodate oxidation. Low concentrations of sodium metaperiodate cause specific oxidative cleavage of the sidechain trihydroxyl system (C_2 - C_3) in sialic acid residues⁶⁻⁸. After the sequential treatment of the erythrocytes with periodate (2 mM, 5 min, 30 °C) and I (5 mM, 2 h, 30 °C), they became highly agglutinable by con A. The lowest lectin concentration, determined by the Microtitre system, which caused visible agglutination of the modified erythrocytes, was 7.5 μ g ml⁻¹ (Table 1). Increasing the length of periodate treatment to 60 min or decreasing it to 1 min did not change this value. That the agglutination is due to the introduction of α -D-mannosyl groups is

Table 1 Lowest lectin concentration which causes agglutination of native and modified bovine erythrocytes

Periodate (mM)	Mannosyl-S-(CH ₂) ₆ - CONH-NH ₂ (mM)	O (CH ₃) ₃ C-O-C-NH-NH ₂ (mM)	Lowest concentration of con A to cause agglutination (μ g ml ⁻¹)
—	—	—	> 500
2	—	—	> 500
2	—	5	> 500
—	5	—	> 500
0.01	5	—	> 500
0.05	5	—	> 500
0.1	5	—	60
0.5	5	—	7.5
1	5	—	7.5
2	5	—	7.5

Washed, packed erythrocytes were incubated with the indicated concentration of sodium metaperiodate for 5 min at 30 °C, washed several times with phosphate-buffered saline (PBS), pH 7.4, and then incubated for 2 h at 30 °C with an equal volume of the appropriate hydrazide (10 mM) or PBS. Packed erythrocytes were also incubated with I or PBS without previous periodate treatment. The Microtitre system (Cooke Engineering) was used to determine the lowest concentration of con A (Sigma grade IV) which caused visible agglutination. Typically, 50 μ l of erythrocytes ($\sim 2.5 \times 10^8$ cells ml⁻¹ in PBS) were incubated with 50 μ l of serially diluted con A solutions in U-type plates. Unperturbed plates were scored after 2 h at room temperature. The concentration of the con A stock solution in this experiment and in all other experiments was calculated from its absorption at 280 nm ($A_{280} = 1.3$ per mg con A ml⁻¹ cm⁻¹).

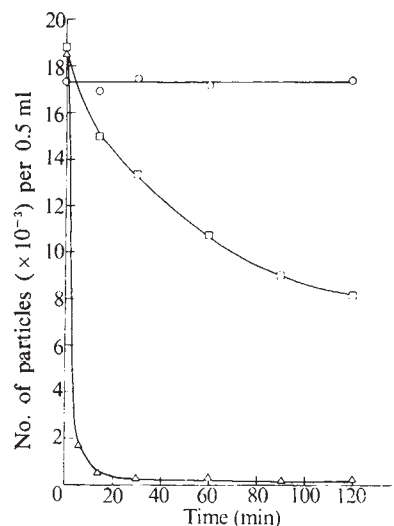
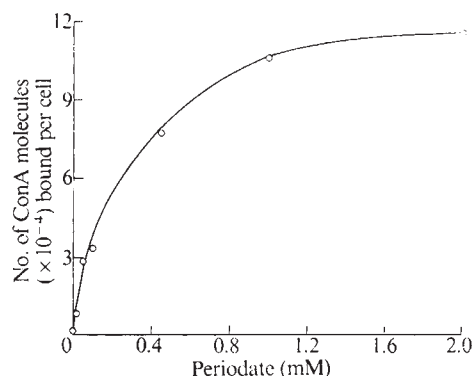


Fig. 1 Kinetics of the con A-induced agglutination of modified erythrocytes. The kinetics of agglutination were studied by following the disappearance of single particles from solution. To 1 ml of con A solution (120 μ g ml⁻¹) was added 50 μ l of the appropriate modified erythrocytes ($\sim 1.5 \times 10^8$ cells ml⁻¹; $\circ$, 0.05 mM; $\square$, 0.1 mM; $\triangle$, 0.5 mM periodate followed by I (5 mM)), and the suspension was incubated at 30 °C on a Dubnoff metabolic shaker. At time intervals, 5 μ l aliquots were removed and added to 10.0 ml of PBS, and the number of single particles determined on a Coulter Counter. The counter was equipped with a 100 μ m aperture and a 0.5-ml manometer. Single erythrocytes were counted at settings of 1/amplification = 1 and 1/aperture current = 1/8, with window setting 20–100. Duplicate samples were removed, each counted in triplicate and the results averaged.

evident from Table 1; periodate treatment alone, periodate treatment followed by a nonspecific hydrazide, or I in the absence of previous periodate treatment did not render the cells agglutinable by con A.

To determine the relationship between the number of aldehyde groups formed and the susceptibility to con A agglutination, the concentration of periodate was varied, whilst keeping the concentrations of I constant (Table 1). With periodate concentrations of 0, 0.01 and 0.05 mM, no agglutination was observed. At 0.5, 1 and 2 mM periodate, the lowest lectin concentration to cause agglutination was 7.5 μ g ml⁻¹. However, at 0.1 mM periodate this value was

Fig. 2 Binding of ¹²⁵I-con A to the modified erythrocytes. Modified erythrocytes were prepared as described in Table 1. 100 μ l of ¹²⁵I-con A (480 μ g ml⁻¹, 7858 c.p.m. μ g⁻¹) was incubated with 100 μ l of erythrocytes ($\sim 7.5 \times 10^8$ cells ml⁻¹) at 30 °C for 40 min. The reaction was stopped by the addition of 4 ml of cold PBS and the suspension centrifuged at 2,000g. The cells were washed a second time with PBS. Both cells and washings were counted in a γ -counter. Experiments were carried out in triplicate and results averaged. Blanks for the correction of nonspecific binding contained 0.1 M α -methylmannoside. The number of erythrocytes per ml in each of the stock solutions was determined accurately with a Coulter Counter, using the settings described in Fig. 1.



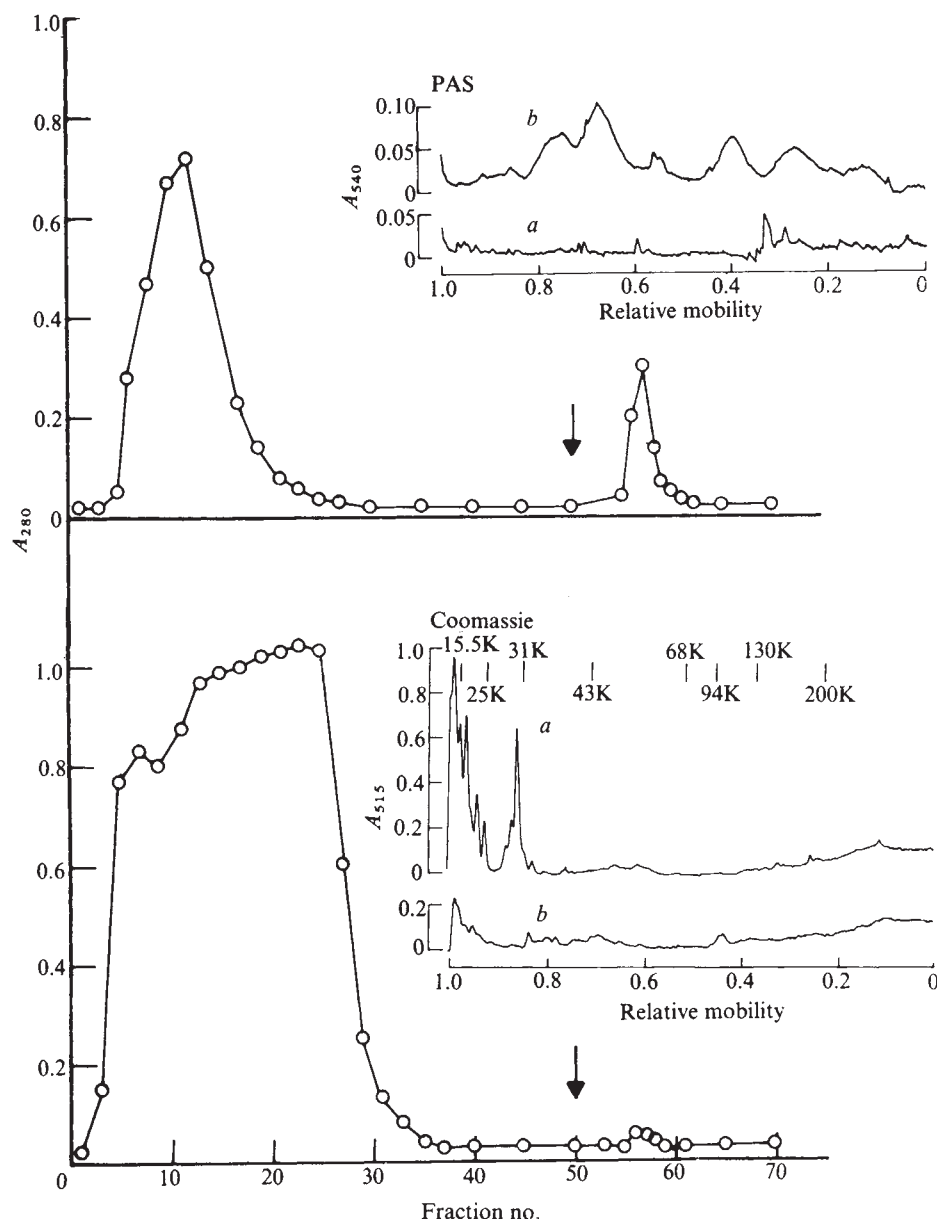


Fig. 3 Isolation of the con A-binding proteins from native and modified bovine erythrocytes. Ghosts were prepared from native and modified erythrocytes (2 mM periodate, 5 min, 30 °C; 5 mM I, 2 h, 30 °C) by the method of Dodge *et al.*¹², and solubilised in 10 mM Tris (pH 7.9) containing 2% deoxycholate. The solubilised membrane proteins were diluted with 10 mM Tris (pH 7.9) to give a final deoxycholate concentration of 1%. Protein samples (43 mg native; 15 mg modified) were applied to a con A-Sepharose 4B column (Sigma; 10 ml; 10 mg protein per ml) and the column eluted with approximately 4 column volumes of 10 mM Tris (pH 7.9) containing 150 mM NaCl. Specifically bound proteins were eluted with 10 mM Tris (pH 7.9) containing 0.1 M α -methylmannoside and 1% deoxycholate (indicated by arrow). The eluate (native, lower; modified, upper) was monitored for protein by absorbance at 280 nm. The various protein fractions were pooled and the protein concentrations determined by a modification of the Lowry method¹³. Protein samples were concentrated by precipitation with 0.1 volumes of 2% acetic acid, the precipitates extracted three times by the addition of ethanol (95%), and the residues dissolved in 0.05 M Tris (pH 6.8) containing 1% SDS. Inserts: gradient polyacrylamide slab gel electrophoresis in 0.1% SDS was performed as described elsewhere¹⁴, with marker proteins in peripheral slots. The marker polypeptides were: myosin (molecular weight 200,000), β -galactosidase (130,000), phosphorylase A (94,000), bovine albumin (68,000), ovalbumin (43,000), DNase (31,000), concanavalin A (25,000), and haemoglobin (15,500). A 3.5% polyacrylamide stacking gel and a 4% to 16% polyacrylamide gradient gel were used. Bromothymol blue was used as the tracking dye and gels were run at 50 V for 16 h. Protein bands (a, native, 16 μ g; b, modified, 14 μ g) were revealed by Coomassie brilliant blue R, glycoproteins located using the periodic acid-Schiff (PAS) procedure¹⁸, and the gels were scanned using a Schoeffel SD3000 densitometer.

raised to 60 μ g ml⁻¹. The kinetics of the con A-induced agglutination were followed using a Coulter counter to observe the disappearance of single cells from solution. Figure 1 shows the effect of con A (120 μ g ml⁻¹) on the kinetics of agglutination of 0.05, 0.1 and 0.5 mM periodate-treated cells. At the highest concentration of periodate the agglutination was extremely rapid; very few single cells were left after 10 min. As expected, the 0.05 mM periodate-treated cells showed no disappearance of single cells, whereas the 0.1 mM periodate-treated cells showed a decrease of approximately 50% after 120 min.

The number of con A-specific binding sites was determined using ¹²⁵I-labelled con A. The binding of ¹²⁵I-con A to the native bovine erythrocytes was characterised by a relatively slow approach to the equilibrium position (approximately 5 h at 30 °C) in contrast to the modified erythrocytes, which were saturated with ¹²⁵I-con A after 30 min at 30 °C (data not shown). Figure 2 shows the binding of ¹²⁵I-con A to the modified erythrocytes. The number of con A molecules bound per cell after 0.05 mM and

0.1 mM periodate treatment were 2.78×10^4 and 3.36×10^4 , respectively. It is apparent, therefore, that the binding of approximately 6,000 additional lectin molecules changes the erythrocytes from being non-agglutinable to agglutinable. By way of comparison, the native bovine erythrocytes bound 1.2×10^5 con A molecules (data not shown) and yet remained non-agglutinable.

To show that the modification had in fact resulted in the formation of new mannose-containing glycoproteins, the con A-binding proteins of native and modified cells were isolated by con A-Sepharose 4B affinity chromatography. The elution profile of native and modified erythrocyte (2 mM periodate, 5 min, 30 °C; 5 mM I, 2 h, 30 °C) membrane proteins are shown in Fig. 3. The amount of specifically eluted protein in each case was 2.2% and 16.3% of the total protein loaded on to the column. The overall recovery of protein in the unabsorbed and specifically eluted fractions was 96% for native erythrocytes and 87% for the modified erythrocytes. The con A receptor from human erythrocytes has been shown to be a component of Band

III (refs 9–11). With the native bovine erythrocytes rather than one specific binding protein, there were several with molecular weights ranging from 15,000 to 30,000. The native bovine receptors do not seem to be sialoglycoproteins, as judged by the absence of any PAS-staining material. However, PAS staining of the modified bovine receptors revealed a radically different profile and showed the presence of at least four major peaks with molecular weights in the region of 3.9×10^4 , 5×10^4 , 1.25×10^5 and 1.8×10^5 .

The mechanism of lectin-induced erythrocyte agglutination is complex. In addition to involving the chemical nature of the receptor itself, it may involve the three-dimensional arrangement of the receptors in the membrane, their mobility within the membrane, and possibly cell surface change. With bovine erythrocytes, it is obvious that the total number of con A binding sites is not correlated with the extent of con A-induced agglutinability. In addition, the net cell surface charge also does not seem to be important, as the simple addition of the mannose residues to the cell surface, without changing net surface charge, rendered the erythrocytes agglutinable. The question then becomes, how are the new receptors different from those already present? It is clear from Fig. 3 that the mannose groups have been added to proteins which are not con A receptors in the native erythrocytes. The possibility exists that these new receptors are more mobile than the old receptors, or that their topological arrangement is such that agglutination is favoured. Experiments are now in progress to distinguish between these possibilities. As previously mentioned, the addition of 6,000 con A binding sites, created at low concentration of periodate, induced the erythrocytes to agglutinate. We would be most interested to know whether these 6,000 sites are introduced on to a unique protein, and, furthermore, the steric relationship of the new receptors to neighbouring proteins and to the lipid bilayer.

Our results establish that it is possible in a specific chemical way to modify the antigenicity of a cell by the introduction of simple sugars on to the cell surface, and also suggest that mannose is the primary specificity determinant of con A-induced agglutination of erythrocytes. This approach could be used generally to introduce defined saccharides into diverse cells in order to modify existing modes of cell-cell interaction. We expect that this technique will lead to a better understanding of the role of the cell surface glycoconjugates in cellular recognition phenomena.

We thank Dr P. Blumberg (Department of Pharmacology, Harvard Medical School) for ^{125}I -concanavalin A and for the use of his Coulter counter.

GEORGE A. ORR
ROBERT R. RANDO

Department of Pharmacology,
Harvard Medical School,
Boston, Massachusetts 02115

Received 21 November 1977; accepted 17 February 1978.

- Gartner, T. K. & Podleski, T. R. *Biochem. biophys. Res. Commun.* **67**, 972–978 (1975); **70**, 1142–1149 (1976).
- Ofek, I., Mirelman, D. & Sharon, N. *Nature* **265**, 623–625 (1977).
- Sumner, J. B. & Howell, S. F. J. *Bact.* **32**, 227–237 (1936).
- Weltzien, H. U. *Expl Cell Res.* **93**, 111–121 (1975).
- Matta, K. L., Girotra, R. N. & Barlow, J. J. *Carbohydr. Res.* **43**, 101–109 (1975).
- Blumenfeld, O. O., Gallop, P. M. & Liao, T. H. *Biochem. biophys. Res. Commun.* **48**, 242–251 (1972).
- Liao, T. H., Gallop, P. M. & Blumenfeld, O. O. *J. biol. Chem.* **248**, 8247–8253 (1973).
- Gahmberg, C. G. & Anderson, L. C. *J. biol. Chem.* **252**, 5888–5894 (1977).
- Findlay, J. B. C. *J. biol. Chem.* **249**, 4398–4403 (1974).
- Barrat, D. G., Sharom, F. J., Thede, A. E. & Grant, C. W. M. *Biochim. biophys. Acta* **465**, 191–197 (1977).
- Roses, A. D. *J. supramolec. Struct.* **4**, 481–486 (1976).
- Dodge, J. T., Mitchell, C. & Hanahan, D. J. *Archs Biochem. Biophys.* **110**, 119–130 (1963).
- Maddy, A. H. & Spooner, R. L. *Vox Sang.* **18**, 34–41 (1970).
- Margolis, J. & Kendrick, K. G. *Analyt. Biochem.* **25**, 347–362 (1968).
- Glossman, H. & Neville, D. M. *J. biol. Chem.* **246**, 6339–6346 (1971).

Possible transcriptional control of three polypeptides which accumulate in a temperature-sensitive mammalian cell line

THE availability of well characterised conditional lethal mutants will widen our knowledge of the genetics and physiology of mammalian cells. Temperature-sensitive mutants (*ts* mutants) have been very useful in studies carried out in a great variety of systems from microorganisms to *Drosophila*. We have undertaken the characterisation of a *ts* mutant (K12) isolated by Roscoe *et al.*¹ from the established line of Chinese hamster fibroblasts Wg1A. The K12 mutant has two interesting properties: it is a cell-cycle mutant^{2,3}, having its execution step (the time during which the mutated function is required by the cells in order to progress through the cell cycle) during the 4 h preceding the initiation of the S phase; and the synthesis of three specific cellular polypeptides is enhanced when the cells are incubated at the non-permissive temperature (40.5 °C) (J.A.M. and V. Fincham, unpublished). These polypeptides, called A, B and C, are not affected in other *ts* mutants defective for the initiation of the S phase and in hybrids in which the K12 mutation is complemented. These data suggest that the K12 mutant is affected in some function necessary both for the commitment of the cells to new rounds of DNA replication and for the regulation of A, B and C polypeptide synthesis. We now present evidence, obtained by investigating the effect of actinomycin D on the proteins synthesised by K12 cells, incubated at the non-permissive temperature and by synthesising the proteins *in vitro*, that the synthesis of these proteins is regulated at the transcriptional level.

Exponentially growing cultures were labelled for 6 h with ^3H -leucine, either in the presence or absence of $5 \mu\text{g ml}^{-1}$ actinomycin D, at 35 °C and 40.5 °C. The concentration of actinomycin D used was enough to decrease the total incorporation of ^3H -uridine into acid-insoluble material to 5% of its normal value in 30 min, whereas it only inhibited incorporation of ^3H -leucine into proteins by 25% during the 6-h period of labelling. At the end of the labelling period, cell extracts were prepared and the proteins analysed by electrophoresis in a polyacrylamide gel containing SDS. Figure 1 shows densitometer tracings of an autoradiograph of the dried gel.

When K12 cells were labelled at 35 °C there were no appreciable differences between the pattern of polypeptides synthesised in the presence or absence of actinomycin D. This indicates that at the permissive temperature the drug does not affect the synthesis of proteins in a specific way. When K12 cells were labelled at 40.5 °C in the absence of actinomycin D, there was an increase in bands A, B and C, corroborating our previous results (J.A.M. and V. Fincham, unpublished). In contrast, when K12 cells were labelled at 40.5 °C, but in the presence of the drug, there was no increase in bands A, B and C. The results of these experiments suggest that the alterations observed in A, B and C polypeptide synthesis are dependent on the continuous production of RNA when K12 cells are incubated at the non-permissive temperature, and indicate that the step at which the K12 mutation regulates the synthesis of these polypeptides is at the transcriptional level.

To provide further evidence supporting this hypothesis, we isolated the polyA-containing RNA from K12 cells incubated at either 35 °C or 40.5 °C and translated it using different cell-free systems. Cultures of cells were incubated at both temperatures for 22 h; cells were collected and lysed; and after removing the nuclei by low speed centrifugation, RNA was extracted from the cytoplasmic fraction. The crude RNA preparation was passed twice through a column of polyU-Sepharose, and the material bound to the column eluted and precipitated with ethanol. This polyA-containing RNA was used in the subsequent cell-free translation experiments.

When polyA-containing RNA, extracted from K12 cells incubated at 40.5 °C, was added to a wheat-germ cell-free system, three specific proteins were observed which were not produced in response to polyA-RNA isolated from K12 cells incubated at 35 °C. The three polypeptides (Fig. 2) had the same apparent molecular weights as bands A, B and C, observed in the *in vivo* experiments. Although the difference in the amount of band A synthesised *in vitro* is less striking than that observed in cells, we conclude from this experiment that qualitatively we are able to reproduce *in vitro* the same change in the synthesis of polypeptides A, B and C that we obtained by labelling the cells *in vivo*.

To rule out the possibility that K12 cells produce mRNAs that are thermosensitive in their translation, we used the ascites cell-free system, which can be incubated *in vitro* at 35 °C and 40.5 °C, the temperatures used in the *in vivo* experiments. The ascites cell-free system was primed with the polyA-containing RNAs extracted from K12 cells incubated either at 35 °C or 40.5 °C and incubated *in vitro* at both temperatures. The proteins synthesised were analysed by electrophoresis in polyacrylamide gels; the results are shown in Fig. 3. The high background, due to endogenous RNA present in this system, obscured the results for bands A and C. However, Fig. 3 shows that band B is produced when the RNA extracted from K12 cells incubated at 40.5 °C is added to the cell-free system, irrespective

Fig. 1 Effect of actinomycin D on the proteins synthesised by K12 cells. Cultures of *ts* mutant K12 cells were grown at 35 °C in Dulbecco's medium⁴, supplemented with 10% calf serum (DC10 medium). When they reached 50% confluency, the medium was replaced by DC10 containing 10% of the normal concentration of leucine and 10 μ Ci ml⁻¹ L-(4,5)-³H-leucine (Amersham, 60 Ci mmol⁻¹); some of the cultures also received 5 μ g ml⁻¹ actinomycin D. At the time of medium change, half of the cultures were shifted to 40.5 °C. Six hours later the cells were scraped off the dishes and washed with PBS; they were resuspended in extraction buffer (0.04 M Tris-HCl pH 8.1, 1 mM 2-mercaptoethanol, 0.2 mM EDTA, 0.01 M MgCl₂ and 0.58 mM phenylmethylsulphonyl chloride) and disrupted by ultrasonication. The extracts were centrifuged at 20,000g for 10 min and the supernatants dialysed against buffer 0.1 M Tris-HCl (pH 6.8) and 10% glycerol. Aliquots were made up 2% in SDS and 0.1 M in DTT; the same number of c.p.m. (about 10⁶) were applied to different slots of the same slab gel containing 5% polyacrylamide in the presence of SDS. The electrophoretic system was as described by Laemmli⁵, and the gels were processed for fluorography⁶. The autoradiograms were scanned using a Joyce-Loebl densitometer. K12 cells labelled at 35 °C in the absence (a) and presence (b) of actinomycin D. K12 cells labelled at 40.5 °C in the absence (c) and presence (d) of the drug.

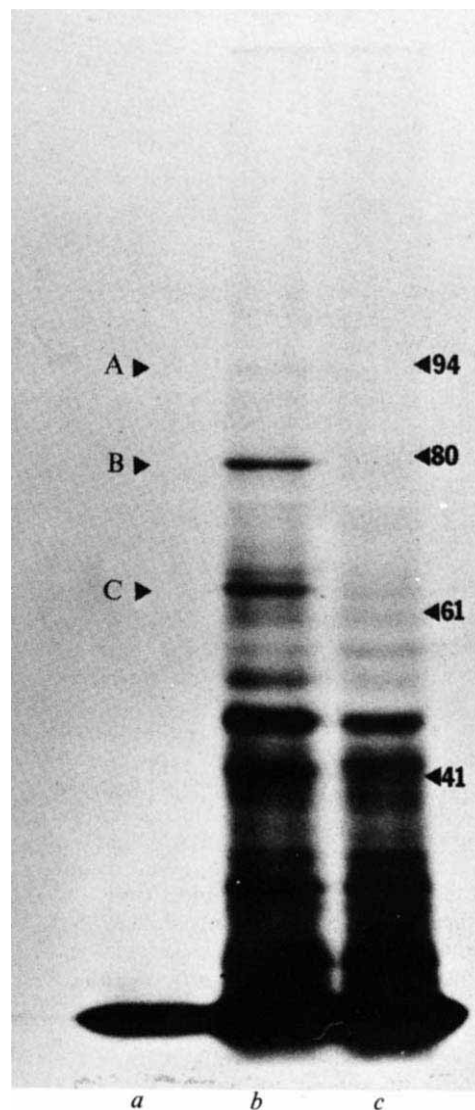
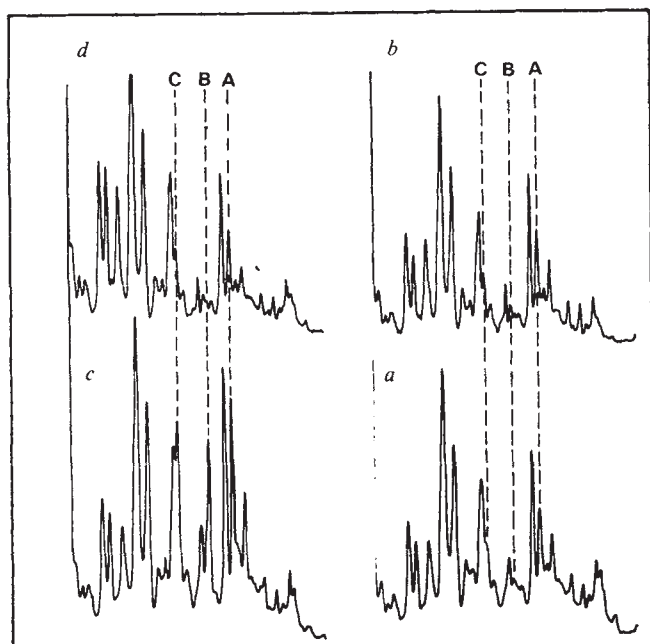


Fig. 2 Translation *in vitro* of the RNAs extracted from K12 cells using the wheat germ system. K12 cells growing exponentially in DC10 medium were split in two halves and incubated at 35 °C and 40.5 °C respectively for 22 h. The cytoplasmic RNA was isolated and the polyA-containing RNA was separated by passage through a column of polyU-Sepharose, as described previously⁷. About 2 μ g polyA-containing RNA were translated using a wheat germ cell-free system⁷, but using ³H-leucine instead of ³⁵S-methionine to label the *in vitro* products. The cell-free products were analysed by electrophoresis on a 7.5% SDS-polyacrylamide slab gel and prepared for fluorography. a, No added mRNA; b, polyA-containing RNA extracted from cells incubated at 40.5 °C; c, and at 35 °C.

of the temperature of incubation *in vitro*. The experiments using the cell-free systems suggest that the mRNAs which code for A, B and C polypeptides accumulate in K12 cells incubated at the non-permissive temperature. Preliminary experiments in which cytoplasmic polyA-containing RNA was separated in a sucrose gradient containing 90% formamide, indicate that each of the polypeptides under study is coded by a separable and therefore different species of mRNA.

Our results do not rule out the possibility of a specific modification of the mRNAs during culture of K12 cells at the high temperature, which activates the mRNAs and makes them more efficient at both temperatures. However, this possibility seems unlikely considering the results obtained with actinomycin D, and together these experiments suggest that the function mutated in K12 cells normally regulates the synthesis of A, B and C polypeptides by limiting the transcription of their genes. An interesting question remains unsolved. How is

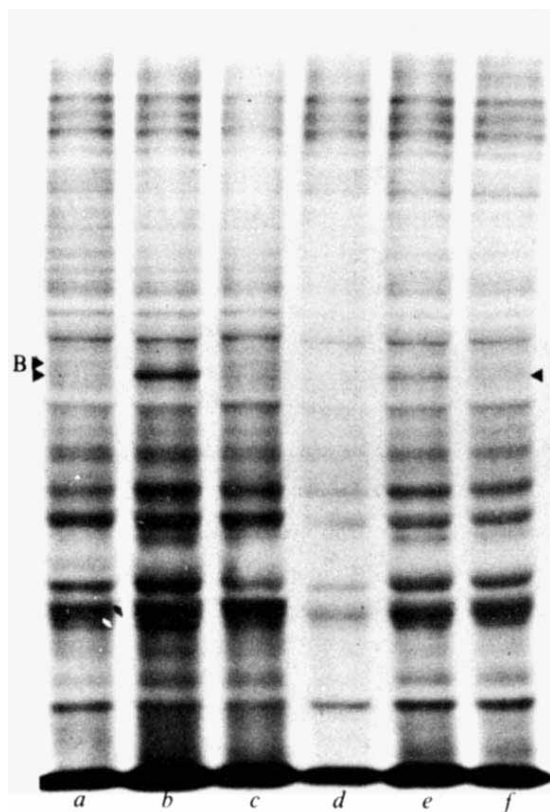


Fig. 3 Translation *in vitro* of the RNAs extracted from K12 cells using the ascites cell-free system: 2 μ g polyA-containing RNAs (described in Fig. 2) were added to a mouse ascites cell-free system⁸. The reactions were carried out either at 35 °C (*a*, *b* and *c*) or 40.5 °C (*d*, *e* and *f*). *a* And *d*, products of the endogenous reactions without RNA added. *b* And *e*, reactions primed by polyA-containing RNA extracted from K12 cells incubated at 40.5 °C. *c* And *f*, reactions primed by RNA extracted from cells incubated at 35 °C.

the transcription of these three polypeptides coordinated? One possibility is that the genes coding for them form an operon-like unit, the transcription of which is regulated by the same factor which is thermosensitive in K12 cells.

If our conclusion is correct, this would be the first report of a mammalian cell in which a cell-cycle mutant is affected in the transcription of specific genes. The isolation and characterisation of more mutants of this kind should widen our knowledge of the mechanisms which regulate the progression of mammalian cells through the cell cycle.

We thank Tricia Wheeler and Valerie Fincham for technical assistance. J.A.M. was supported by an EMBO Fellowship.

JOSE A. MELERO*
ALAN E. SMITH

Tumour Virology and Translation Laboratories,
Imperial Cancer Research Fund,
Lincoln's Inn Fields, PO Box 123,
London WC2, UK*

Received 21 December 1977; accepted 22 February 1978.

*Present address: New York University Medical Center, Department of Pathology, 550 First Avenue, New York 1006.

1. Roscoe, D. H., Read, M. & Robinson, H. *J. cell. Physiol.* **82**, 325 (1973).
2. Roscoe, D. H., Robinson, H. & Carbonell, A. W. *J. cell. Physiol.* **82**, 333 (1973).
3. Smith, B. J. & Wigglesworth, N. W. *J. cell. Physiol.* **82**, 339 (1973).
4. Dulbecco, R. & Freeman, G. *Virology* **8**, 396 (1959).
5. Laemmli, U. K. *Nature* **227**, 680 (1970).
6. Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56**, 335 (1975).
7. Wheeler, T., Bayley, S. T., Harvey, R., Crawford, L. V. & Smith, A. E. *J. Virol.* **21**, 215 (1977).
8. Pawson, T., Martin, G. S. & Smith, A. E. *J. Virol.* **19**, 950 (1976).

Activity of an individual serotonergic neurone in *Aplysia* enhances synthesis of cyclic adenosine monophosphate

BRIEF stimulation of the marine mollusc *Aplysia* by food elicits an arousal response that is in many ways analogous to arousal seen in vertebrates. Animals assume a characteristic 'orientating' posture, and the speed and strength of successive biting responses progressively increase¹. In addition, food arousal is associated with an acceleration of heart rate² and modification of defensive reflexes³. Several lines of evidence suggest that arousal in vertebrates may be mediated by biogenic amines, acting to alter neuronal excitability through an effect involving stimulation of cyclic adenosine monophosphate (cAMP)⁴⁻⁶. Previous studies have indicated that activity of a pair of identified serotonergic cells (the metacerebral cell, MCC) in *Aplysia* and other gastropod molluscs can produce both peripheral and central effects associated with modifications of the biting response during food arousal⁶⁻¹¹. Thus, for example, in *Aplysia*, firing of an MCC enhances the contractility of the accessory radular closer muscles, muscles involved in the biting response; MCC firing also increases the frequency of the rhythmic burst output from the buccal ganglion which controls the biting and swallowing phases of feeding^{6,11}. The effects of the MCC seem to be due to the release of serotonin at its central and peripheral terminals. Because of the actions of serotonin in mediating aspects of arousal of feeding behaviour^{6,12} and perhaps sensitisation of other reflexes in *Aplysia*^{11,14,18}, it has been suggested that serotonin may act in a way analogous to that of epinephrine and other biogenic amines in the central and peripheral effects associated with the arousal of vertebrates^{4,12,13}. Possible evidence for the involvement of the biogenic amines in arousal and sensitisation is the ability of these compounds to stimulate the synthesis of cAMP and thereby to produce relatively long-lasting effects in excitable tissue. Several observations are consistent with the hypothesis that the action of the MCC at buccal muscle may involve a second messenger, such as cAMP^{11,15}; these are that the onset and decay of the effects of the MCC are remarkably slow for synaptic transmission, that the action of the MCC and serotonin in buccal muscle seems to involve an unusual neurotransmitter mechanism in which there is no alteration in the resting membrane potential, and that phosphodiesterase-resistant analogues of cAMP simulate the action of the MCC on buccal muscle. We provide evidence here that food arousal in *Aplysia* may also involve the action of a biogenic amine acting to increase cAMP levels, and we show that the application of serotonin or the firing of an identified serotonergic neurone enhances the synthesis of cAMP in muscles involved in the biting response.

To measure the formation of cAMP from a radioactive precursor, individual accessory radular closer (ARC) muscles were incubated for 110 min in 100 μ l of 3.3×10^{-5} M ³H-adenosine (30 Ci mmol⁻¹) in artificial seawater. After they were rinsed, the muscles were placed for 5 min in 0.9 ml of seawater containing serotonin at concentrations of 10^{-3} – 10^{-10} M. The stimulation was stopped by homogenising the muscles in a solution of ice cold 6% trichloroacetic acid, with ¹⁴C-cAMP added to monitor for recovery of cAMP. The homogenate was centrifuged at 2,000 r.p.m., and the supernatant was processed for cAMP using the sequential chromatography method of Salomon *et al.*¹⁶.

Figure 1 shows the dose-response curve for serotonin stimulation of cAMP. The concentration of serotonin needed to stimulate synthesis of cAMP half maximally was 3×10^{-6} M, while maximal stimulation of synthesis, which was approximately 200 times greater than control, occurred

at 10^{-4} M serotonin. Acetylcholine in concentrations up to 10^{-3} M had no effects on cAMP synthesis. A threefold stimulation of cAMP occurred when muscles were stimulated with solutions containing 10^{-8} M serotonin. This value corresponds to the threshold concentration of serotonin needed to produce potentiation of muscle contractions in *Aplysia*¹¹.

Having found that serotonin produced a marked stimulation of cAMP synthesis, we investigated whether this was accompanied by an increase in the total amount of cAMP in the muscle; this was assayed by the isotopic displacement method of Gillman¹⁷. The results obtained showed that as the rate of cAMP synthesis was stimulated by serotonin, total levels of cAMP in the muscle increased in a dose-related manner. Because exogenous serotonin, the transmitter of the MCC, was capable of stimulating the synthesis of cAMP, we next investigated the possibility that firing of the MCC could release sufficient serotonin to produce measurable effects on the synthesis of cAMP in the muscle. As the MCC in *Aplysia* innervates only the ipsilateral ARC muscle, for these experiments the muscle contralateral to the stimulated MCC was used as a control. The left and right ARC muscles, together with the cerebral and buccal ganglia, were dissected as previously described⁶, and the muscles were placed in a small subchamber. An MCC and an ipsilateral motor neurone were both impaled, and the continuity of the axon of the MCC with its corresponding muscle was verified by determining the ability of the MCC to produce potentiation of muscle contractions evoked by stimulation of the motor neurone⁶. The electrodes penetrating the MCC and motor neurone were withdrawn, and the solution bathing the muscle was replaced with 0.6 ml of 6.67×10^{-5} M ^3H -adenosine (30 Ci mmol^{-1}) in seawater. After 4 h, the solution of adenosine was removed and the muscles were rinsed. The MCC was reimpaled and stimulated by intracellular depolarising current. The mean duration of stimulation was 40 s, the mean number of spikes per train was 297 and the mean frequency of stimulation $7.6 \text{ spikes s}^{-1}$. At the end of stimulation, both muscles were frozen with liquid nitrogen and were processed as in the previous experiment, with the exception that the extraction was performed by homogenising the muscles in 80% ethanol.

Figure 2 shows the results of five such experiments. In each experiment, newly synthesised cAMP was higher in the muscle on the side in which the MCC was fired than in the control muscle. The mean amount of labelled cAMP was 3.5 times higher in the experimental muscles ($t = 4.17$; d.f. = 4, $P < 0.015$; two-tailed t -test for correlated means). In three additional experiments the effects of phosphodiesterase digestion on the product of MCC stimulation was evaluated. These experiments confirmed the ability of MCC to stimulate cAMP; the mean amount of newly synthesised cAMP

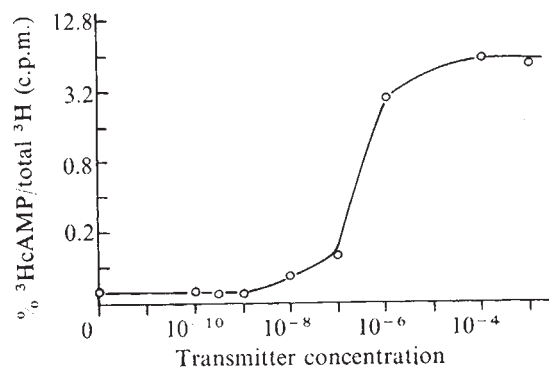


Fig. 1 Dose-response curve for serotonin stimulation of cAMP synthesis. Synthesis was expressed as % of ^3H -cAMP over the total ^3H radioactivity.

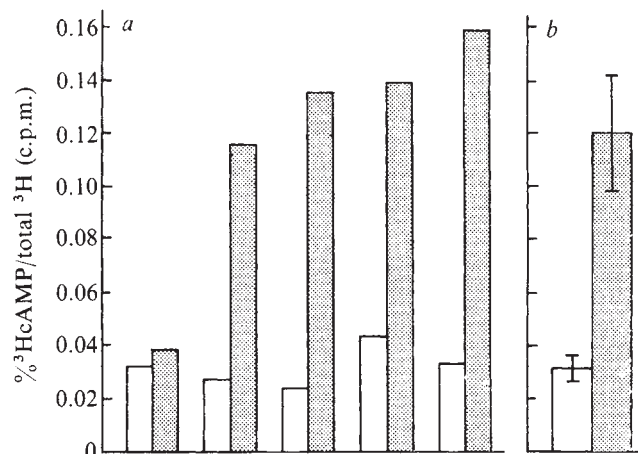


Fig. 2 Stimulation of cAMP synthesis by activation of the MCC. Muscles contralateral to the stimulated side were used as control. a, Results of five individual experiments; b, mean of the 5 experiments shown on the left. Vertical bars indicate standard errors. □, control; ▨, stimulated.

in the muscle ipsilateral to the stimulated side was 5.5 times higher than in the control muscle. The product of phosphodiesterase digestion co-migrated chromatographically with the breakdown product of ^{14}C -cAMP. This provides additional evidence that the product of MCC stimulation identified as ^3H -cAMP was indeed *bona fide* cAMP.

In previous studies of the role of neurotransmitters in stimulating synthesis of cAMP, synthesis was enhanced by electrically stimulating groups of neurones, or by applying transmitters to the bathing solutions. The present findings indicate that cAMP synthesis can be enhanced by activity of a single identifiable neurone, utilising a known neurotransmitter. In *Aplysia* buccal muscles, as in the vertebrate heart, the action of a biogenic amine in enhancing contraction may utilise cAMP as a second messenger. These results further strengthen the hypothesis that arousal in vertebrates and invertebrates may involve certain common molecular mechanisms^{11-13,15,18}.

We thank Drs E. Kandel and J. Schwartz for their comments on an earlier draft of this manuscript, and K. Hilten for figures. This work was supported in part by NINCDS Grant NS 12492, Grant 1 PO1 GM 23540 Scope D, Alfred P. Sloan Foundation Training Grant and NIH medical scientist training grant GM 02042.

KLAUDIUSZ R. WEISS
MICHAEL SCHONBERG
DAVID E. MANDELBAUM
IRVING KUPFERMANN

Division of Neurobiology and Behavior,
Departments of Physiology, Psychiatry and Neurology,
College of Physicians and Surgeons of Columbia University,
and The New York State Psychiatric Institute,
New York, New York 10032

Received 9 December 1977; accepted 13 February 1978.

- Susswein, A. J., Weiss, K. R. & Kupfermann, I. *J. comp. Physiol.* (in the press).
- Dieringer, N., Koester, J. & Weiss, K. R. *J. comp. Physiol.* (in the press).
- Advokat, C., Carew, T. & Kandel, E. *Abstr. Soc. Neurosci.* 2, 313 (1976).
- Bolme, P., Fuxe, K. & Lidbrink, P. *Res. Commun. chem. path. Pharmac.* 4, 657-697 (1972).
- Mansour, T. E. & Stone, D. B. *Biochem. Pharmac.* 19, 1137-1146 (1970).
- Weiss, K. R., Cohen, J. & Kupfermann, I. *Brain Res.* 99, 381-386 (1975).
- Berry, M. S. & Pentreath, V. W. *J. exp. Biol.* 65, 361-380 (1976).
- Gillette, R. & Davis, W. J. *J. comp. Physiol.* 116, 129-159 (1977).
- Granzow, B. & Kater, S. B. *Neuroscience* (in the press).
- Prior, D. J. & Gelperin, A. *J. comp. Physiol.* 114, 217-232 (1977).
- Weiss, K. R., Cohen, J. L. & Kupfermann, I. *J. Neurophysiol.* 41, 181-203 (1978).
- Weiss, K. R. & Kupfermann, I. *Soc. Neurosci. Symp.* 3, (in the press).
- Kandel, E. R., Brunelli, M., Byrne, J. & Castellucci, V. *Cold Spring Harb. Symp. quant. Biol.* 40, 465-482 (1976).
- Mayeri, E., Koester, J., Kupfermann, I., Liebeswar, G. & Kandel, E. R. *J. Neurophysiol.* 37, 458-475 (1974).
- Weiss, K. R., Schonberg, M., Cohen, J., Mandelbaum, D. & Kupfermann, I. *Abstr. Soc. Neurosci.* 2, 338 (1976).
- Salomon, Y., Londos, C. & Rodbell, M. *Analyt. Biochem.* 58, 541-548 (1974).
- Gillman, A. G. *Proc. natn. Acad. Sci. U.S.A.* 67, 305-309 (1970).
- Brunelli, M., Castellucci, V. & Kandel, E. R. *Science* 194, 1178-1180 (1976).

Effects of the synthetic enkephalin analogue FK 33-824 in man

THE synthetic analogue of methionine enkephalin, FK 33-824, exerts a long-lasting and potent analgesic effect in animals after systemic and enteral administration¹. In high doses it produces tolerance, cross-tolerance with morphine, and dependence in monkeys. Like β -endorphin, FK 33-824 produces akinesia in rats and rabbits. All effects seen in animals can be abolished by naloxone. We report here on the results of the first studies performed with FK 33-824 in man (tolerance, electroencephalogram (EEG), endocrine profile and plasma levels), as a basis for future therapeutic studies.

Forty healthy male volunteers (age 20–38 yr) who were fully informed of the investigational character of the studies and who gave written informed consent, received single intramuscular (i.m.) injections of FK 33-824 in the dose range of 0.1–1.2 mg. Six subjects received 0.5 mg twice, two subjects were given the drug by slow intravenous (i.v.) infusion (0.5 mg in 30 min) and five subjects received intradermal injections (0.01 mg). FK 33-824 was available in lyophilised form which was diluted with physiological saline. The studies were performed open or single-blind, in most cases with randomised placebo controls.

Blood pressure and pulse rate (supine and standing), respiratory rate and body temperature were followed up to 6 h after drug administration and showed no relevant changes. ECG (30 min, 2 h) and standard haematological and clinical chemistry tests (24 h) were not altered. However, subjective and objective side effects were registered. The most striking symptom, and one that occurred in every subject, was a feeling of 'heaviness' in all the muscles of the body, often combined with a feeling of oppression on the chest or tightness in the throat which induced a certain amount of anxiety. These symptoms were reported 3–5 min after the injection of FK 33-824, reached peak intensity over 5–15 min and faded away within 15–30 min.

Other symptoms occurring in 50–60% of the subjects were an impressive increase in bowel sounds (time of onset 3–10 min, duration 30–60 min), redness of the face, injection of conjunctivae and chemosis (time of onset 5–30 min, duration 1–3 h). The same symptoms occurred shortly after starting a slow intravenous infusion, when a cumulative dose of 0.025 mg had been given. The side effects increased only slightly in intensity when the infusion was continued up to 0.5 mg. In two subjects, a flush of the whole body was observed and in one subject an attack of rhinitis vasomotorica occurred 15 min after the i.m. injection of 0.5 mg FK 33-824. Classical symptoms of morphine², such as changes in emotional behaviour or mental alertness, formication and nausea were not observed. When injected intradermally in five subjects, 0.01 mg FK 33-824 provoked a flare reaction with a mean area of 8.9 cm² (s.d. $\pm$ 4.0) which was reduced to 4.7 cm² (s.d. $\pm$ 4.4) by oral pre-

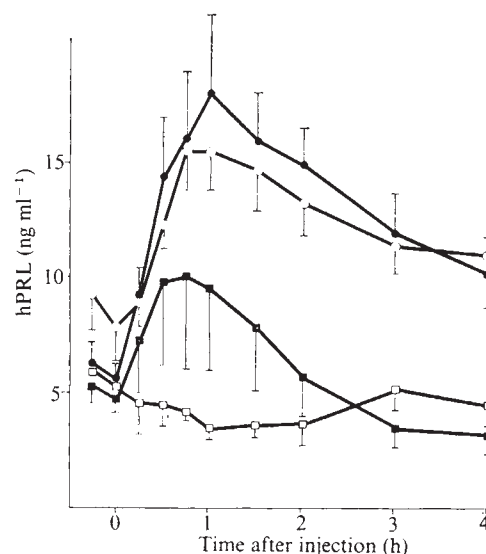


Fig. 1 Mean plasma levels ($\pm$ s.e.m.) of hPRL in normal subjects after single intramuscular doses of FK 33-824. ●, 1.0 mg ($n = 8$); ○, 0.5 mg ($n = 6$); ■, 0.1–0.2 mg ($n = 4$); □, saline ($n = 4$).

treatment (interval 5 h) with 2 mg of ketotifen, a drug possessing antianaphylactic and antihistaminic activity ($P < 0.05$ compared to placebo).

A dose-dependent, stimulatory effect on the secretion of prolactin (hPRL) and growth hormone (hGH) was observed after single i.m. injections of FK 33-824 (Figs 1 and 2). Peak levels were reached after 45–60 min, and 4 h after drug administration, hPRL-values were still elevated, whereas the effect on hGH had subsided within 2–3 h. Similar results have been reported with morphine in rats^{3,4} and man⁵, and with β -endorphin in rats⁶. However, contrary to results obtained with β -endorphin⁶ and naloxone, pre-treatment with the morphine antagonist nalorphine (10 mg i.m. 15 min before FK 33-824) did not abolish the stimulatory effect of FK 33-824 on hPRL secretion (Table 1). Stress was probably not a reason for the changes in hPRL and hGH, because in six subjects who showed marked increases in hPRL- and hGH-levels, the plasma cortisol levels remained within normal limits.

In 12 subjects, the EEG profile was examined after single i.m. doses of FK 33-824 (0.1–1.2 mg). Spectral analysis demonstrated a slowing of α waves, an increase of total electrical output in the 1–40 Hz-band and an increase in the β band up to a maximum of 50%. A decrease in slow waves was seen regularly. The increase of total electrical output was considered to be due to the increase of the α band, seen by visual evaluation as slowing, hypersynchrony and spindling. Maximum changes occurred 2–4 h after administration. These effects were not reversed by nalorphine (10 mg i.m. 90 min after FK 33-824). They are similar to those provoked by opioid analgesics only with respect to the action on the α frequency⁷.

Table 1 Effects of pretreatment with nalorphine on FK 33-824-induced stimulation of hPRL secretion

Subject	–15 min	0	hPRL (ng ml ⁻¹)					
			15 min	30 min	45 min	60 min	90 min	120 min
M.J. a	14.0	13.0	12.5	19.0	21.0	18.9	17.0	17.4
b	9.7	9.2	12.1	39.4	53.0	54.8	51.0	39.8
H.T. a	12.3	6.4	7.8	11.3	15.9	17.0	10.0	8.0
b	6.3	6.5	6.1	9.8	15.1	15.9	16.8	12.5
E.U. a	5.4	4.6	6.6	12.9	19.2	20.5	20.8	15.0
b	7.4	6.2	6.0	12.9	19.5	22.9	37.0	34.4

a, 0.5 mg FK 33-824 i.m. (time 0).

b, 10 mg nalorphine i.m. 15 min before 0.5 mg FK 33-824 i.m. (time 0).

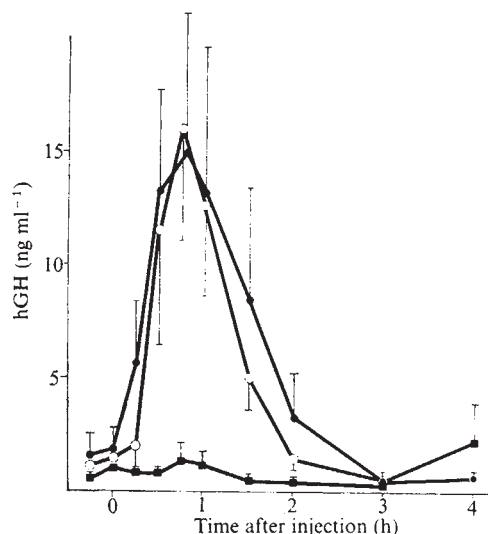


Fig. 2 Mean plasma levels ($\pm$ s.e.m.) of hGH in normal subjects after single intramuscular doses of FK 33-824. ●, 1.0 mg ($n = 7$); ○, 0.5 mg ($n = 6$); ■, 0.1–0.2 mg ($n = 4$).

Plasma levels of FK 33-824 were followed in several subjects receiving different doses (Table 2). Dose-dependent peak levels were measured at 30 min, with only a slight decline recorded over the following 2 h.

Results of the comparative animal studies with FK 33-824 and morphine¹ suggested that we would find morphine-like effects in man with single parenteral doses between 2 and 10 mg. Not only were the doses found to be effective in normal subjects very much lower (0.025–1.2 mg), but the clinical effects observed were unexpected, not morphine-like and not prevented by nalorphine (10 mg). The EEG changes and the stimulatory effect on secretion of hPRL were also not reversed by nalorphine. It is therefore doubtful that all of these effects are mediated through stimulation of 'morphine receptors'. We have so far not been able to clarify the mechanism of these effects. An influence of the galenical form used in man is highly unlikely, as there were no qualitative or quantitative differences between the pharmacological properties of the lyophilised and the pure drug.

Plasma lactate and pyruvate and electromyograph recordings were unchanged when examined at the time of maximal muscular symptoms. The rapid onset of 'anaphylactoid' vascular and gastrointestinal symptoms suggests either an unspecific 'peptide' effect or the release of endogenous humoral factors acting on the smooth musculature of small vessels and the gut (for example, histamine, serotonin). However, results of experiments performed to clarify this hypothesis were negative; although the cutaneous flare reaction was suppressed by ketotifen, the clinical symptoms

Table 2 Plasma concentrations of FK 33-824 in ng ml⁻¹ $\pm$ s.e.m. after single intramuscular injections

Dose <i>n</i>	0.1 mg 1	0.2 mg 1	0.5 mg 4	1.0 mg 2
15 min	—	—	—	30.1 $\pm$ 5.3
30 min	16	33	26.3 $\pm$ 1.6	51.7 $\pm$ 10.8
60 min	15	25	26.1 $\pm$ 0.7	29.4 $\pm$ 2.8
90 min	—	—	24.1 $\pm$ 2.2	24.4 $\pm$ 2.3
120 min	8	14	21.2 $\pm$ 0.9	—

Plasma levels of FK 33-824 were determined by means of a radio-receptor assay (D. Hauser and A. Closs, in preparation). Specific binding of ³H-naloxone (1 nM) to rat brain membranes was measured in the presence of plasma (treated and untreated subjects). A standard displacement curve for FK 33-824 was determined with control plasma, thus considering the inhibition of ³H-naloxone binding by plasma alone. The lower limit of sensitivity was 2–5 ng ml⁻¹.

were not changed by pretreatment with the histamine-1-receptor blocking drugs clemastine (2 mg i.v. 10 min before FK 33-824) and ketotifen (1 mg twice daily orally for 2 d), or with the histamine-2-receptor blocking agent cimetidine (200 mg four times daily for 2 d), each tested in one subject. The clinical effects observed are therefore not mediated by stimulation of histamine receptors alone. The release of serotonin is probably also not involved, as in eight subjects, plasma levels of serotonin were unchanged 30 min after 0.5 mg of FK 33-824 i.m.; moreover, no increased release of serotonin from thrombocytes collected 5 min after injection of the same dose was observed in three subjects, and the clinical effects were not prevented by pretreatment with methysergide (3 mg twice daily orally for 2 d) in one subject. Also, atropine (0.5 mg i.v.) did not influence them when injected 10 min after 0.5 mg FK 33-824 i.m.. Further investigations related to the possible release of other endogenous substances or peptides, such as vasopressin or those of the paracrine system of the gut, where enkephalin immunoreactivity has been demonstrated⁸, are needed. An ultimate explanation for the non-morphine-like type of effects seen might be the presence of enkephalin receptors in other tissues than those now known, with a specificity differing from that of morphine. On the other hand, the fact that affinity to rat opiate receptors was demonstrated in the plasma using a radioreceptor assay (Table 2) indicates that a loss of the postulated biological activity by metabolic or other processes can be excluded and that the expected effects of FK 33-824 might occur in higher doses than those so far studied.

BEAT VON GRAFFENRIED
EMILIO DEL POZO
JIRI ROUBICEK

Department of Experimental Therapeutics,
Biological and Medical Research Division, Sandoz Ltd,
CH-4002 Basle, Switzerland

EVA KREBS
WALTER PÖLDINGER

Psychiatric Clinic Wil,
CH-9500 Wil, Switzerland

PETER BURMEISTER
LOTHAR KERP

Medical University Clinic,
D-78 Freiburg i.Br., FRG

Received 5 December 1977; accepted 2 March 1978.

1. Roemer, D. *et al.* *Nature* **268**, 547–549 (1977).
2. Smith, G. M. & Beecher, H. K. *J. Pharm. exp. Ther.* **136**, 47–52 (1962).
3. Brown, M. & Vale, W. *Endocrinology* **97**, 1151–1156 (1975).
4. McCann, S. M. *et al.* in *Narcotics and the Hypothalamus* (eds Zimmermann, E. & Hecht, H.) 121–134 (Raven, New York, 1974).
5. Tolis, G. *et al.* *J. clin. Endocr. Metab.* **41**, 797–800 (1975).
6. Rivier, C. *et al.* *Endocrinology* **100**, 238–241 (1977).
7. Fink, M. *et al.* in *Narcotic Drugs* (ed. Clouet, D.) 452–466 (Plenum, New York, 1971).
8. Polak, J. M. *et al.* *Lancet* **i**, 972–974 (1977).

Polyamine deficiency reduces the rate of DNA replication fork movement in *Escherichia coli*

STUDIES with microbial mutants^{1–3} and animal cells^{4,5} have shown that spermidine and spermine, and perhaps their biosynthetic precursor putrescine, are required for optimal cell proliferation. Mutants of *Escherichia coli* blocked in the synthesis of putrescine provide an important tool with which to study the cellular function of these compounds. Previous work suggested a defect in DNA replication during starvation for polyamines; infection of spermidine-starved *E. coli* with bacteriophage T4 resulted in a submaximal rate of DNA accumulation⁶, while polyamine-limited growth of *E. coli* resulted in about twice as much DNA per cell as expected from the growth rate². Also, increased cellular DNA has been seen during thymine-

limited chromosome replication⁷ and in strains possessing mutations in the *rep* locus⁸. In both cases, this change in cellular composition seemed to be due to a reduced rate of replication fork movement and an increased number of forks per cell^{9,10}. We show here that the abnormally high DNA content during polyamine limitation also seems to be the result of decreased replication fork velocity.

The DNA content of strain DK6 (ref. 2) was measured during growth in the presence and absence of added polyamines (see the legend to Fig. 1 for experimental details). Both spermidine-starved and unstarved cells (generation times of 54 and 30 min, respectively) contain about 21 μg DNA per 10^9 cells. This is atypical behaviour, as *E. coli* generally exhibits a decreased DNA content with decreasing growth rate¹². For example, DK6 has a generation time of 52 min when the supplemental amino acids are omitted from the otherwise complete culture medium. These cells, which now grow prototrophically at a reduced rate identical to that of the spermidine-deficient cells, contain 9 μg DNA per 10^9 cells as compared with the value of 21 μg for the starved, spermidine-deficient cells.

Dividing DNA content by generation time yields average rates of DNA accumulation in starved and unstarved cells of 0.39 μg per 10^9 cells per min and 0.70 μg per 10^9 cells per min, respectively. The decreased rate during polyamine starvation could be due to a lowered frequency of initiation, a decreased rate of growing fork movement, or the presence of a gap in the DNA synthetic cycle. To distinguish between these alternatives, we have measured the number of DNA strands being elongated in starved and unstarved cells. This value was then compared with the net rate of DNA synthesis.

The number of elongating strands was measured using the approach of Lark and Bird¹¹. Cellular DNA was pulse labelled with ³H-thymidine and the labelled chromosomes were subsequently allowed to segregate in non-radioactive medium for several generations. The linear portion of the plot of log % labelled cells (determined autoradiographically) against the number of generations in non-radioactive medium should extrapolate back to the number of strands originally labelled by the pulse. Figure 1 shows the segregation of labelled chromosomes in DK6 growing in the presence and absence of putrescine. Points from both starved and unstarved cultures fall along the same line, which extrapolates back to an average of about five labelled strands per cell. Thus, although the overall rate of DNA replication was reduced during polyamine starvation, the number of strands of replicating DNA was the same in both starved and unstarved cells. All of the starved and unstarved cells were labelled by the original ³H-thymidine pulse, indicating that there is little or no gap in the DNA synthetic cycle in either condition.

An independent estimate of the number of rounds of DNA synthesis in progress (n), and the replication time (C), can be obtained by measuring ΔG , the increment in the amount of DNA after inhibiting protein synthesis^{9,18}. Blocking protein synthesis allows rounds in progress to finish, but prevents new rounds from initiating. Thus, the more rounds in progress the larger ΔG will be. Knowing ΔG allows one to calculate n , and C can be calculated from n and the generation time. Results from this type of experiment are given in Table 1. Starved and unstarved DK6 have similar values for n , in agreement with the autoradiographic data presented above. These values of n translate into C values of 44 min and 80 min for unstarved and starved cells, respectively. As a control, this type of experiment was also performed on DK5, the parent strain of DK6. DK5 does not require putrescine, and its growth rate was manipulated by varying the growth medium to produce generation times similar to those of starved and unstarved DK6. In this case one sees the expected difference in the n values and the expected similarity in the C

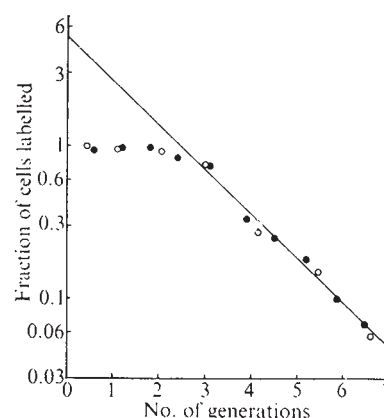


Fig. 1 Segregation of pulse labelled chromosomes in starved and unstarved DK6. Cells of DK6 (ref. 2) were grown on MOPS medium¹⁴ at 37 °C in a rotary shaker. The medium was supplemented with 0.2% glucose and a mixture of amino acids¹⁵, as noted in the text. Cells were starved as previously described² or were supplemented with putrescine (100 $\mu\text{g ml}^{-1}$). Growth was followed turbidimetrically at 450 nm. Cells growing in the presence (○) or absence (●) of putrescine were labelled for 10% of their generation time with ³H-thymidine (3.2 $\mu\text{g ml}^{-1}$, 28 $\mu\text{Ci } \mu\text{g}^{-1}$). The cells were then filtered and resuspended in pre-warmed, non-radioactive medium at time zero and samples were periodically removed for autoradiographic determination of the fraction of labelled cells. Autoradiography was carried out according to the method of Caro and Van Tubergen¹⁶ using Kodak NTB2 emulsion. Cells were stained with 0.02% crystal violet in 30% ethanol and viewed by phase contrast microscopy. The percentage of labelled cells was determined according to Caro¹⁷. The number of generations spent in non-radioactive medium was determined by turbidity and by counting (in a haemocytometer) cells that had been diluted into 4% formaldehyde during the experiment.

values (Table 1).

Data presented here demonstrate that, with and without putrescine starvation, DK6 has the same number of replication forks and the same DNA content. Because the generation times with and without putrescine differ by a factor of almost 2, it follows that the chain elongation time (C in Table 1) must also differ by a factor of almost 2. This same conclusion was reached independently by measurements of the increase in DNA content following chloramphenicol addition. The velocity of growing fork movement is usually not the major site of regulation of cellular DNA synthesis. The rate of chromosome replication is adjusted to the cellular growth rate by variations in the frequency of initiation¹². As polyamine deficiency results in a decreased rate of fork movement rather than a decreased initiation frequency, it would seem that cellular polyamine levels do not normally regulate DNA replication. It may be that polyamines participate directly as cofactors in the DNA synthetic reaction. However, our experiments do not rule out the possibility that the reduced rate of fork

Table 1 Increment in DNA synthesis after chloramphenicol addition

Strain	Culture conditions	τ	ΔG	n	C
DK6	+ putrescine	29	61.8	1.52	44
	- putrescine	55	58.5	1.45	80
DK5	glucose + amino acids	34	72.1	1.73	59
	glucose	55	36.0	0.94	52

Cells were grown as described in Fig. 1. The generation time (τ) was determined by optical density measurements at 450 nm. ΔG (the % increment in the amount of DNA after blocking protein synthesis) was determined as follows. At zero time chloramphenicol was added to a final concentration of 200 $\mu\text{g ml}^{-1}$. Three samples were taken for DNA determination at zero time with periodic sampling thereafter. DNA content was determined using the diphenylamine reaction described by Burton¹¹ using calf thymus DNA as a standard. The average of the plateau values and the average zero time value were then used to calculate the % increase in the DNA content. This ΔG value was then used to calculate n (the number of rounds in progress per chromosome) and C (the replication time) as described by Pritchard and Zaritsky⁹.

movement is an event which is removed from the primary site(s) of polyamine action.

Is the reduced rate of replication fork movement a direct result of polyamine deprivation? Spermidine has been shown to have a stimulatory effect on several *in vitro* systems that may be related to cellular DNA replication. Conversion of single-stranded phage DNA to its replicative form *in vitro* is stimulated by spermidine¹⁹⁻²². *E. coli* DNA binding protein, which is required for *in vitro* DNA synthesis, will catalyse the reassociation of DNA in physiological conditions in the presence of polyamines²³. DNA gyrase, an enzyme which may control the super-coiling of DNA in *E. coli* during replication^{24,25}, is also stimulated *in vitro* by spermidine²⁴. At present it is impossible to ascertain the relevance, if any, of these *in vitro* interactions of spermidine to the *in vitro* results reported here. However, it is interesting to speculate that if gyrase activity *in vivo* were to be reduced by polyamine limitation, the resulting abnormalities in chromosome configuration could account for the reduced rate of RNA chain elongation²⁶ as well as the defect in DNA replication fork movement.

In addition to these results with *E. coli*, studies in mammalian cell systems also suggest a role for polyamines in DNA synthesis. Drug-induced polyamine deficiency in activated lymphocytes and in HeLa cells resulted in a reduced rate of ³H-thymidine incorporation, both in intact cells^{5,27} and in nuclei isolated from these cells^{27,28}. However in contrast to our results with *E. coli*, the lowered ³H-thymidine incorporation in HeLa cells seems to be due to a lowered percentage of active replicons, with only a small effect on elongation²⁷.

This work was supported in part by grants from the US NSF (BMS 75-05381), the American Cancer Society (NP 195) and the NIH (GM 13957). L. G. was supported by NIH Training Grant GM 00052.

LEE E. GEIGER
DAVID R. MORRIS

Department of Biochemistry,
University of Washington,
Seattle, Washington 98195

Received 12 December 1977; accepted 27 February 1978.

1. Maas, W. K., Leifer, Z. & Poindexter, J. *Ann. N.Y. Acad. Sci.* **171**, 957-967 (1970).
2. Morris, D. R. & Jorstad, C. J. *Bact.* **113**, 271-277 (1973).
3. Whitney, P. & Morris, D. R. *J. Bact.* (in the press).
4. Fillingame, R. H., Jorstad, C. M. & Morris, D. R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4042-4045 (1975).
5. Mamont, P. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **73**, 1626-2630 (1976).
6. Dion, A. S. & Cohen, S. S. *J. Virol.* **9**, 423-427 (1972).
7. Zaritsky, A. & Pritchard, R. H. *J. Bact.* **114**, 824-837 (1973).
8. Lane, H. E. D. & Denhardt, D. T. *J. Bact.* **120**, 805-814 (1974).
9. Pritchard, R. H. & Zaritsky, A. *Nature* **226**, 126-131 (1970).
10. Lane, H. E. D. & Denhardt, D. T. *J. molec. Biol.* **97**, 99-112 (1975).
11. Burton, K. *Biochem. J.* **62**, 315-323 (1956).
12. Cooper, S. & Helmstetter, C. E. *J. molec. Biol.* **31**, 519-540 (1968).
13. Lark, K. G. & Bird, R. E. *Proc. natn. Acad. Sci. U.S.A.* **54**, 1444-1450 (1965).
14. Neidhardt, F. C., Block, P. L. & Smith, D. F. *J. Bact.* **119**, 736-747 (1974).
15. Maas, W. K. *Molec. Gen. Genet.* **119**, 1-9 (1972).
16. Caro, L. G. & van Tubergen, R. P. *J. Cell Biol.* **15**, 173-188 (1962).
17. Caro, L. G. *J. molec. Biol.* **48**, 329-338 (1970).
18. Sueoka, N. & Yoshikawa, H. *Genetics* **52**, 747-757 (1965).
19. Scheckman, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **69**, 2691-2695 (1972).
20. Scheckman, R., Weiner, A. & Kornberg, A. *Science* **186**, 987-993 (1974).
21. Geider, K. & Kornberg, A. *J. biol. Chem.* **249**, 3999-4005 (1974).
22. Weiner, J. H., Bertsch, L. & Kornberg, A. *J. biol. Chem.* **250**, 1972-1980 (1975).
23. Christiansen, C. & Baldwin, R. L. *J. molec. Biol.* **115**, 441-454 (1977).
24. Gellert, M., Mizuuchi, K., O'Dea, M. & Nash, H. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3872-3876 (1976).
25. Gellert, M., O'Dea, M., Itoh, T. & Tomizawa, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4474-4478 (1976).
26. Morris, D. R. & Hansen, M. T. *J. Bact.* **116**, 588-592 (1973).
27. Krokan, H. & Eriksen, A. *Eur. J. Biochem.* **72**, 501-508 (1977).
28. Knutson, J. C. & Morris, D. R. *Biochim. biophys. Acta* (in the press).

Mutations in *nirA* gene of *Aspergillus nidulans* and nitrogen metabolism

In both prokaryotes and eukaryotes many structural genes are subject to more than one form of control. The ways in which these various forms of regulation interact with each other are therefore of considerable importance for an understanding of the control of gene expression. In the ascomycete fungus *Aspergillus nidulans*, the structural

genes for nitrate reductase and nitrite reductase, the enzymes of nitrate assimilation, are under the control of two positive regulatory genes, *nirA* (refs 1-3) and *areA* (ref. 4). The *nirA* gene product is required for the synthesis of nitrate and nitrite reductases and mediates nitrate induction of these enzymes. The *areA* gene product is necessary for the synthesis of a wide range of permeases and enzymes involved in nitrogen metabolism, including nitrate and nitrite reductases, and mediates ammonium repression of these permeases and enzymes. Previously available data have indicated that both *nirA* and *areA* products would be required for the synthesis of nitrate and nitrite reductases, the role of *nirA* being largely confined to nitrate induction and that of *areA* being confined to ammonium repression. Here we show that a more complex situation exists. A new allele of the *nirA* gene, designated *nirA*^d-101 and derived from the wild-type allele in two mutational steps, results in both constitutivity and derepression and eliminates the need for a functional *areA* product.

Two kinds of mutation in the *nirA* gene have been described: *nirA*⁻ mutations (designated *nirB* in the earliest publications), the loss of function class, lead to non-inducibility and a consequent inability to utilise nitrate and nitrite as nitrogen sources^{1-3,5,6}. One *nirA*^c mutation (*nirA*^c-1) has been characterised¹⁻³. It leads to constitutive synthesis of nitrate reductase and semi-constitutive synthesis of nitrite reductase, but the synthesis of these enzymes is still strongly ammonium-repressible. Because so many structural genes are under *areA* control, mutations in the *areA* gene are highly pleiotropic and quite heterogeneous^{4,7}. The most common class of mutation, designated *areA*^r, leads to an inability to utilise nitrogen sources (including nitrate and nitrite) other than ammonium and to a correspondingly dramatic reduction in the levels of many permeases and enzymes. The *areA*^r phenotype is that of a loss of function mutation: *areA*^r-18 does not revert intracistronically and has not proved separable from a translocation breakpoint⁸. Another class of *areA* mutation, designated *areA*^d, leads to derepression of one or more ammonium-repressible activities⁴. However, although *areA*^d mutations can result in derepressed synthesis of nitrate and nitrite reductases in the presence of ammonium (ref. 4 and D. J. Cove and H. N. A., Jr, unpublished), there is no evidence to suggest that any *areA* allele can result in constitutive synthesis of these enzymes; nitrate induction mediated by a functional *nirA* product is required.

The highly pleiotropic phenotype of an *areA*^r mutation affords a powerful positive selection procedure for obtaining a variety of other regulatory mutations. These result in the recovery of the ability to utilise one or more nitrogen sources. The use of this procedure to select new *nirA* alleles was suggested by the finding that *nirA*^c-1 suppresses certain *areA*^r mutations which are leaky (for example, *areA*^r-120) or temperature-sensitive (for example, *areA*^r-130) for utilisation of nitrite and, in most cases, nitrate as well. More extreme *areA*^r alleles such as *areA*^r-1⁴ and *areA*^r-18 (the total loss of function allele) are not suppressed by *nirA*^c-1. Without postulating an interaction between the *nirA* and *areA* products, there are at least two ways in which *nirA*^c mutations might suppress certain *areA*^r alleles for utilisation of nitrate and nitrite. If uptake of nitrate and nitrite be under *areA* control, then uptake of these ions as inducers could be the limiting factor for their utilisation by particular *areA*^r mutants. Thus *nirA*^c mutations could compensate for inducer exclusion. NADPH has been shown to prevent inactivation of at least nitrate reductase⁹. Increased stability of this enzyme in *nirA*^c-1 strains is correlated with elevated levels of both NADPH and the NADPH-generating enzymes, glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase^{9,10}. *nirA*^c mutations

Table 1 Nitrate and nitrite reductase activities of wild-type and *nirA* mutant strains

Relevant genotype	Nitrogen source(s)	Relative activities	
		Nitrate reductase	Nitrite reductase
Wild-type (<i>nirA</i> ⁺)	NO ₃ ⁻	100	100
	Urea	6	2
	NO ₃ ⁻ + NH ₄ ⁺	2	9
	NH ₄ ⁺	2	2
<i>nirA</i> ^c -1	NO ₃ ⁻	98	117
	Urea	75	29
	NO ₃ ⁻ + NH ₄ ⁺	20	21
	NH ₄ ⁺	3	6
<i>nirA</i> ^c -200	NO ₃ ⁻	155	140
	Urea	45	13
	NO ₃ ⁻ + NH ₄ ⁺	16	20
	NH ₄ ⁺	6	4
<i>nirA</i> ^c -201	NO ₃ ⁻	94	79
	Urea	111	40
	NO ₃ ⁻ + NH ₄ ⁺	5	7
	NH ₄ ⁺	13	9
<i>nirA</i> ^c -202	NO ₃ ⁻	83	120
	Urea	107	32
	NO ₃ ⁻ + NH ₄ ⁺	8	13
	NH ₄ ⁺	11	7
<i>nirA</i> ^c -203	NO ₃ ⁻	94	90
	Urea	101	33
	NO ₃ ⁻ + NH ₄ ⁺	8	13
	NH ₄ ⁺	8	7
<i>nirA</i> -300	NO ₃ ⁻	95	86
	Urea	5	4
	NO ₃ ⁻ + NH ₄ ⁺	11	13
<i>nirA</i> ^d -101	NO ₃ ⁻	143	111
	Urea	85	27
	NO ₃ ⁻ + NH ₄ ⁺	79	63
	NH ₄ ⁺	85	40

The *nirA*^c-1, *nirA*-300 and *nirA*^d-101 strains carry *biA*-1 (biotin auxotrophy), whereas all other strains carry *pabaA*-1 (*p*-aminobenzoate auxotrophy). All strains carry the *areA*⁺ allele. Mycelia were grown for 20 h at 25 °C in minimal medium¹⁶ containing 1% (w/v) D-glucose as carbon source and the nitrogen sources indicated above at 10 mM (ammonium as the (+)-tartrate, nitrate as the sodium salt) or 5 mM (urea). Growth conditions, extraction procedures and determinations of soluble protein, nitrate reductase (EC 1.6.6.3) and nitrite reductase (EC 1.6.6.4) have been described previously⁸. Relative enzyme activities are expressed as percentages of the specific activities of the wild-type strain grown on nitrate as nitrogen source. The induction and selection of *nirA*^c-1 has been described previously¹. Other *nirA* mutations described here were induced using ultraviolet light and selected as follows: *nirA*^c-200 was obtained by replica plating¹⁷ nitrate-utilising revertants of a strain of genotype *pabaA*-1 *areA*^c-120 on to glucose-minimal medium containing 10 mM NH₄⁺ and 100 mM ClO₃⁻. Although *areA*^c-120 is leaky, it apparently results in ammonium derepressed synthesis of those activities it partially retains (possibly because of loss of an ammonium-repressible ammonium permease⁸). Thus we found that *areA*^c-120 *nirA*^c-1 double mutants are sensitive to 100 mM ClO₃⁻ in the presence of 10 mM NH₄⁺, and the *areA*^c-120 *nirA*^c-200 double mutant was recognised because it has the same phenotype. *nirA*^c-201, -202 and -203 were obtained in the same way as *nirA*^c-200 except that the replica plates contained no chlorate but were instead stained for the presence of nitrite in colonies after a substrate treatment with 1 M NaNO₃ for 1 h at 37 °C. These *areA*^c-120 *nirA*^c-200, -201, -202 and -203 double mutants, like *areA*^c-120 *nirA*^c-1 double mutants, stain pink in these conditions. *nirA*^c-1, -200, -201, -202 and -203 suppress *areA*^c-120 for nitrite as well as nitrate utilisation. *nirA*-300 was selected as a suppressor mutation in a strain of genotype *biA*-1 *pabaA*-1 *areA*^c-130 *fwA*-1 (having fawn conical colour) as allowing nitrite utilisation. It suppresses both *areA*^c-120 and *areA*^c-130 for nitrate and nitrite utilisation. *nirA*^d-101 was selected as a suppressor mutation in a strain of genotype *biA*-1 *pabaA*-1 *areA*^c-18 *nirA*^c-1 as allowing utilisation of nitrite. All of these new *nirA* mutations have been shown to be tightly linked to *nirA*^c-1 or a *nirA*⁻ mutation.

might therefore increase enzyme stability by raising NADPH levels, thereby enhancing the growth of those *areA*⁺ mutants which are able to produce at least some enzyme.

Table 1 shows the effects on enzyme levels of four new *nirA*^c alleles (in *areA*⁺ strains), selected as suppressors of the leaky mutation *areA*^c-120. These mutations, for reasons which are as yet obscure, resemble *nirA*^c-1 in conferring only semi-constitutive synthesis of nitrite reductase. However, *nirA*-300, selected as suppressing the temperature-sensitive mutation *areA*^c-130, does not lead to constitutivity or even to significant ammonium derepression (Table 1). Preliminary evidence (K. N. Rand and H. N. A., Jr, unpublished) suggests that it elevates levels of the NADPH-generating enzymes, which might explain its ability to suppress.

Extensive efforts to revert a strain carrying the total loss of function allele *areA*^c-18 in a *nirA*⁺ background on nitrate or nitrite have failed. However, one revertant was obtained using an *areA*^c-18 *nirA*^c-1 double mutant. This revertant apparently has a new allele of the *nirA* gene, designated *nirA*^d 101. *nirA*^d-101 suppresses the other *areA*^c mutations tested (*areA*^c-1, -120, and -130) as well as *areA*^c-18 for both nitrate and nitrite utilisation, thus indicating locus-specific suppression. *nirA*^d-101 retains the constitutivity characteristic of *nirA*^c-1 but in addition leads to strong ammonium derepression (Table 1). After a preliminary cross had established tight linkage of *nirA*^d-101 to a *nirA*⁻ mutation, 1,045 progeny were analysed from a cross of a *nirA*^d-101 strain to a *nirA*⁺ strain. None carry *nirA*^c-1, *nirA*^d-101 strains are easily distinguishable from *nirA*^c-1 strains, because 10 mM ammonium protects *nirA*^c-1 but not *nirA*^d-101 strains against the toxicity of 100 mM chlorate. Chlorate sensitivity is closely associated with nitrate reductase levels (but its basis is controversial, see ref. 6). The chlorate sensitivity of *nirA*^d-101 strains in the presence of ammonium is a striking confirmation of their ammonium-derepressed synthesis of nitrate reductase. The progeny of the cross of the *nirA*^d-101 strain to the *nirA*⁺ strain were also screened with a view to detecting recombinants carrying the second mutational lesion of *nirA*^d-101 without *nirA*^c-1, the first mutation. None of the progeny show reduced nitrate or nitrite utilisation, and none are derepressed without also being constitutive. This latter phenotype is possibly the most obvious one to expect for strains carrying only the second *nirA*^d-101 mutation, but the second mutation might equally well be silent or completely inseparable from the *nirA*^c-1 mutation. *nirA*^d-101 is partially dominant to both *nirA*^c-1 and *nirA*⁺ in diploids homozygous for *areA*^c-1 with respect to suppression of *areA*^c-1 for utilisation of nitrate and nitrite as nitrogen sources. Partial, rather than complete, dominance results from a dose effect in the case of *nirA*^c-1² and probably therefore also in the case of *nirA*^d-101.

The ability of a mutant *nirA* allele to alleviate the requirement for the *areA* product has a parallel in a prokaryotic system which likewise involves the joint action of two positive regulatory molecules for structural gene expression. Expression of the L-arabinose regulon of *Escherichia coli* is under control of *araC*, a positively and negatively acting regulatory gene mediating induction, and *crp*, the gene specifying the positive acting cyclic AMP receptor protein mediating carbon catabolite repression¹¹⁻¹³. A mutant *araC* allele, designated *araC*¹, allows gene expression in the absence of the cyclic AMP receptor protein and results in insensitivity to carbon catabolite repression^{14,15}. It will be interesting to discover whether the ability of mutations in pathway-specific positive regulatory genes such as *nirA* and *araC* to bypass the requirements for the products of positive regulatory genes having more extensive domains such as *areA* and *crp* is a frequent feature of systems subject to more than one form of positive control.

We thank Hugh Penfold, David Cove, Claudio Scazzocchio, Don MacDonald, Sue Jones, Chris Bailey, David Tollervey and Phil Oliver for helpful comments and Paul Helliwell for technical assistance. We thank the MRC for a studentship (K.N.R.) and the Royal Society for the Smithson Research Fellowship (H.N.A.). This work was partially supported by a S.R.C. grant to Professor J. M. Thoday.

Note added in proof: A report has just appeared¹⁸ showing that certain mutations in the *araC* gene of *E. coli* selected for constitutivity also result in decreased sensitivity to carbon catabolite repression and partial alleviation of the requirement for the cyclic AMP receptor protein for expression of the arabinose operon.

KEITH N. RAND*
HERBERT N. ARST, JR

Department of Genetics,
University of Cambridge,
Downing Street, Cambridge, UK

Received 9 November 1977; accepted 24 February 1978.

*Present address: Technical Services, Glaxo Laboratories Ltd, North Lonsdale Road, Ulverston, Cumbria, UK.

1. Pateman, J. A. & Cove, D. J. *Nature* **215**, 1234–1237 (1967).
2. Cove, D. J. *Nature* **224**, 272–273 (1969).
3. Cove, D. J. *Proc. R. Soc. Lond. B* **176**, 267–275 (1970).
4. Arst, H. N., Jr & Cove, D. J. *Molec. gen. Genet.* **126**, 111–141 (1973).
5. Pateman, J. A., Rever, B. M. & Cove, D. J. *Biochem. J.* **104**, 103–111 (1967).
6. Cove, D. J. *Molec. gen. Genet.* **146**, 147–159 (1976).
7. Arst, H. N., Jr & Scazzocchio, C. *Nature* **254**, 31–34 (1975).
8. Rand, K. N. & Arst, H. N., Jr *Molec. gen. Genet.* **155**, 67–75 (1977).
9. Dunn-Coleman, N. S. & Pateman, J. A. *Molec. gen. Genet.* **152**, 285–293 (1977).
10. Hankinson, O. & Cove, D. J. *J. biol. Chem.* **249**, 2344–2353 (1974).
11. Englesberg, E. & Wilcox, G. A. *Rev. Genet.* **8**, 219–242 (1974).
12. Lee, N., Wilcox, G., Gielow, W., Arnold, J., Cleary, P. & Englesberg, E. *Proc. natn. Acad. Sci. U.S.A.* **71**, 634–638 (1974).
13. Pastan, I. & Adhya, S. *Bact. Rev.* **40**, 527–551 (1976).
14. Heffernan, L., Bass, R. & Englesberg, E. *J. Bact.* **126**, 1119–1131 (1976).
15. Heffernan, L. & Wilcox, G. *J. Bact.* **126**, 1132–1135 (1976).
16. Cove, D. J. *Biochim. biophys. Acta* **113**, 51–56 (1966).
17. Mackintosh, M. E. & Pritchard, R. H. *Genet. Res.* **4**, 320–322 (1963).
18. MacInnes, K. R., Sheppard, D. E. & Falgout, B. *J. Bact.* **133**, 178–184 (1978).

Structural and biological link between pyrethroids and DDT in new insecticides

WE developed previously a structural model of a nerve membrane receptor for DDT-type insecticides¹ which enabled us to synthesise isosteric structures^{2,3} with high insecticidal activity. Concurrently, Elliott and co-workers^{4,5} were developing highly active insecticides based on the modification of the structure of naturally occurring pyrethrum. We noted that some of our most active DDT isosteres (Fig. 1) contained a similar structural feature with the pyrethroids, namely the relatively uncommon di-substituted cyclopropane ring. Recently Van den Bercken⁶ found a close relationship between the action of these two classes of compounds, in experiments with DDT and the pyrethroid allethrin (III, Fig. 1) in the sensory nervous system of the toad (*Xenopus laevis*). This prompted us to examine the synthesis of compounds which would combine in one molecule the structural features of both the DDT and pyrethroid insecticides. We report here the structures and properties of such compounds recently synthesised in this laboratory.

All new compounds and the standards used in this study were evaluated biologically in mortality tests on a susceptible strain of the housefly (*Musca domestica*, WHO/In/1) and the Australian sheep blowfly (*Lucilia cuprina* W.). To obtain the basic insect toxicity without the intervening metabolic detoxication of the compounds, the mortalities for both insects were also measured in the presence of a synergist. The contact repellency values⁷ from attractant targets, treated with sublethal doses of the compounds, were also determined for the two insects. We also compared the neurophysiological effects of selected compounds from each series, for their spontaneous generation of multiple nerve impulses at the labellar chemosensory receptor of the blowfly.

Initially we synthesised the dimethylcyclopropane isostere (II, Fig. 1) of the DDT analogue¹, 1,1-bis (*p*-

ethoxyphenyl)-2,2-dichlorocyclopropane (I, Fig. 1). For the latter we had previously³ measured the multiplicity of induced nerve impulses when applied to the chemosensory receptor of the fly. We found that for DDT itself or its halocyclopropane analogues there was a direct relationship between this activity and mortality in the same strain of the insect.

The dimethylcyclopropane II, unlike its dichlorocyclopropane isostere I, does not yield multiple impulses (Fig. 2) proportional to its considerable insecticidal activity. In this effect it resembles the potent pyrethroid insecticides such as permethrin (Fig. 2), and differs from other members of the DDT series.

This lack of generation of the characteristic trains of nerve impulses can be explained by the findings of Narahashi⁸ for allethrin and Burt⁹ for other pyrethroids—that insect synapses or postsynaptic neurones are the most likely primary targets for pyrethroids. The sensory neurones in our preparation do not synapse and therefore we cannot detect the primary nerve-transmission disturbance for the pyrethroids, in contrast to the DDT-type compounds. This qualitative difference (Fig. 2) in neurophysiological response can be used to distinguish between the two classes of insecticides.

In the subsequent design of the intermediate DDT-

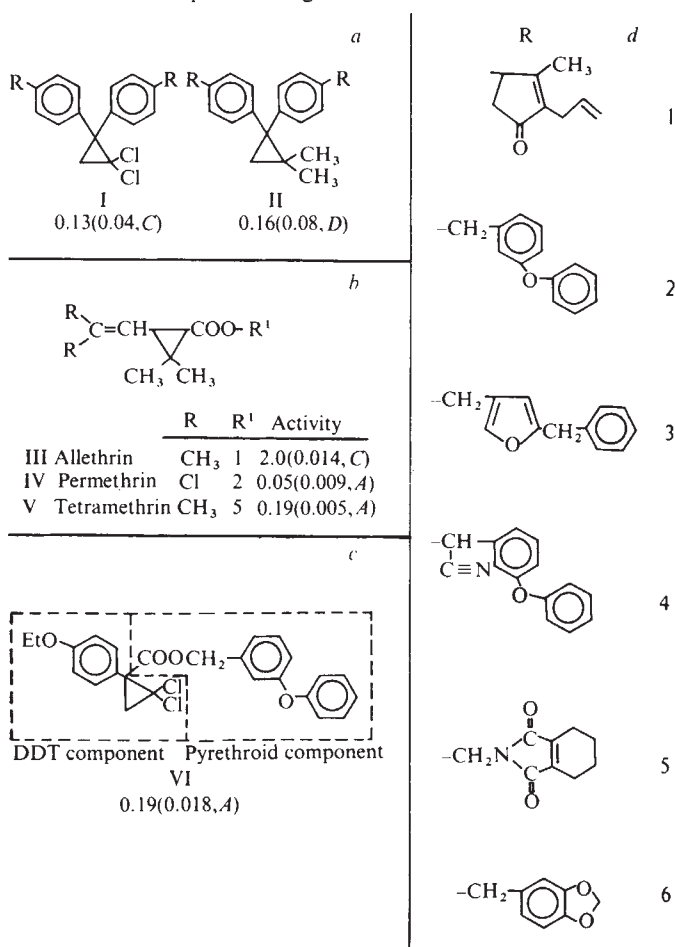


Fig. 1 Comparison of structures and insecticidal activity for a, DDT-isosteres; b, pyrethroids; c, a combined DDT-pyrethroid structure; d, alcohol radicals of natural and synthetic pyrethroids. Mortality values (LD₅₀ µg per ♀ insect for the housefly) are given for each compound. LD₅₀ values for the same compounds synergised with sesamex, 2-(3,4-methylenedioxyphenoxy)3,6,9-trioxadecane, 1 µg applied with each dose of compound in acetone are given in parentheses. The capital letters in parentheses are rankings of repellency measured by residence⁷ of insects on attractant treated targets; D, repellency index value 0–50; C, 50–65; B 65–80; A, 80–100. D represents an inactive compound and B ranking represents activity previously extended to field applications for the control of insect pests¹⁹.

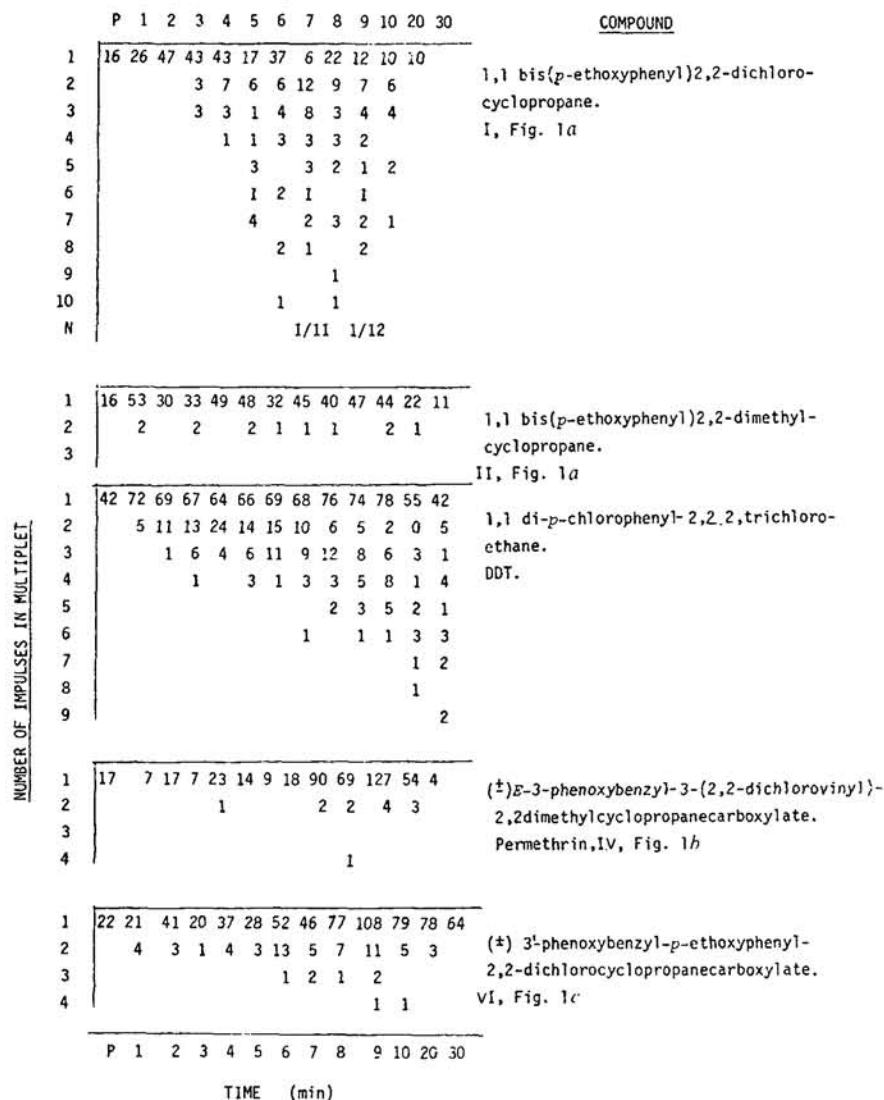


Fig. 2 Histograms of averages of replicates of nerve-impulse multiplicity measured at the labellar-hair chemoreceptor of the blowfly (*Lucilia cuprina* Weid.) A qualitative difference exists between DDT, its dichlorocyclopropane analogue and the dimethylcyclopropane II, the pyrethroid permethrin, and the combined structure VI. The latter group shows a marked decrease in this neurophysiological response. The insecticides (0.001 M in olive oil) were applied to the base of the hair of the labellum of the proboscis only, of the blowfly, impaled on a silver-hair electrode. The active glass-capillary electrode (0.01 M, LiCl) was mounted in a Neurolog preamplifier. —P— are control impulses recorded before treatment. At this and each subsequent time interval impulse activity was recorded for 10 s stored in a time-shift register and the time (± 0.001 s) between each impulse was printed out. Time intervals between impulses < 10 ms (ref. 20) were considered as multiplets.

pyrethroid structures we kept to the requirements of our postulated model² for DDT compounds, encompassing within its parameters the alcohol moiety of the active pyrethroids (Fig. 1d). From this model it could be deduced that the symmetry and size¹ of the insecticides was limited only in one part of the molecule, the 'apex' cyclopropane structure. Additionally it was seen that of the two possible isomers the 1-aryl-1-cyclopropanecarboxylate illustrated in Fig. 1c, was preferred.

Many compounds with the other isomeric structure (1-aryl-3-cyclopropanecarboxylate) have been reported, mainly in patent literature¹⁰⁻¹². Most of the reports contained little or no information on their insecticidal activity.

Where activity was reported, the compounds possessed a fraction of that of their pyrethroid models. To establish the difference in activity between the isomers, we synthesised the 3'-phenoxybenzyl-3-*p*-ethoxyphenyl-2,2-dichlorocyclopropane-1-carboxylate, the isomer of compound VI (Fig. 1c) that does not fit our DDT model. This compound was 50–60-fold less active with or without synergist compared to isomer VI, and did not possess any repellent properties to either test insect. Its dimethylcyclopropane congener, the racemic 'wrong' isomer of compound XVI (Table 1), also displayed a similar drop in insecticidal activity.

The new DDT-pyrethroid combination structures proved to be active insecticides. Of the large number of derivatives synthesised, some relevant to this study are listed in Table 1. Their structural similarity and biological properties relate them to both classes of insecticides.

Several derivatives, in common with the pyrethroids, were

potent repellents of the housefly or the blowfly. The new compounds also induced blowfly sensory-nerve multiplicity patterns qualitatively resembling those of pyrethroids (Fig. 2). The optical resolution of the racemic esters gave an increase in the insecticidal activity for one enantiomer in contrast to DDT derivatives¹³. For a more detailed comparison of this effect, the 1-*p*-ethoxyphenyl-2,2-dichlorocyclopropane-1-carboxylic acid was resolved and the absolute configuration (*R*) of the laevorotatory enantiomer was determined by X-ray crystallography. After esterification, the *R*(-)-3'-phenoxybenzyl ester (VII, Table 1) gave a greater insect mortality than the *S*(+) ester (VIII, Table 1). The increased activity for the *R*-enantiomer does relate our compounds to the pyrethroids³. In our series also¹⁴, the resolution of the dimethylcyclopropane derivative yielded no difference between the activities of the two enantiomeric esters (XVI and XVII, Table 1).

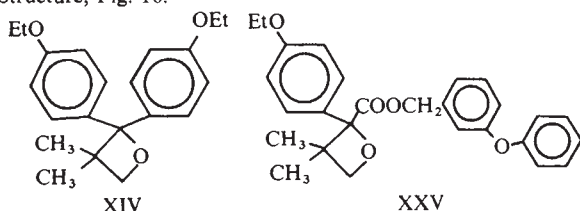
We found that the activities of the esters broadly followed DDT-analogue substitution patterns in the aryl ring and the cyclopropane parts of the acid component, while the alcohol radicals followed the activities for the same components in the pyrethroids. There were, however, several exceptions to this pattern. The difluorocyclopropanes, inactive in the DDT series¹, possess high activity in the combined form (XII, XIII, Table 1). Another exception is the activity of the 3,4-methylenedioxybenzyl ester (XI, Table 1), to our knowledge not previously reported as an active¹⁵ alcohol radical of pyrethroids. On the other hand, the tetrahydrophthalimidomethyl alcohol radical (5, Fig. 1d) which is the component of pyrethroid insecticide

Table 1 Toxicity and repellency values of new insecticides to the housefly (*Musca domestica*) and the blowfly (*Lucilia cuprina*)

No.	R	X,X	Alcohol component (number, structure from Fig. 1d)	LD ₅₀ *		(LD ₅₀ + synergist)*		Repellency*	
				Housefly	Blowfly	Housefly	Blowfly	Housefly	Blowfly
VI	Ethoxy	Dichloro	3-Phenoxybenzyl	(2) 0.19 (0.018)	0.19 (0.021)	A	C		
VII	R (-) Ethoxy	Dichloro	3-Phenoxybenzyl	(2) 0.076 (0.0016)	0.022 (0.002)	B	D		
VIII	S (+) Ethoxy	Dichloro	3-Phenoxybenzyl	(2) > 2 (0.031)		—	D		
IX	Ethoxy	Dichloro	5-Benzyl-3-furymethyl	(3) 0.29 (0.0011)	0.022 (0.0034)	A	C		
X	Ethoxy	Dichloro	α-cyano-3-phenoxybenzyl	(4) 0.22 (0.029)	0.017 (0.0001)	B	A		
XI	Ethoxy	Dichloro	3,4-Methylenedioxybenzyl	(6) 0.32 (0.05)	0.33 (0.0098)	D	D		
XII	Ethoxy	Difluoro	3-Phenoxybenzyl	(2) 0.55 (0.0028)	0.075 (0.0046)	B	B		
XIII	Ethoxy	Difluoro	5-Benzyl-3-furymethyl	(3) 0.32 (0.0095)	0.010 (0.0001)	A	A		
XIV	R (-) Ethoxy	Difluoro	3-Phenoxybenzyl	(2) 0.22 (0.002)	—	A	—		
XV	S (+) Ethoxy	Difluoro	3-Phenoxybenzyl	(2) > 32 (0.041)	—	C	—		
XVI	R (-) Ethoxy	Dimethyl	3-Phenoxybenzyl	(2) 1.98 (0.029)	0.58 (0.031)	D	D		
XVII	S (+) Ethoxy	Dimethyl	3-Phenoxybenzyl	(2) 1.00 (0.023)	0.35 (0.028)	D	D		
XVIII	Ethoxy	Dibromo	3-Phenoxybenzyl	(2) 0.18 (0.015)	0.25 (0.048)	C	D		
XIX	Ethoxy	Bromo, chloro	3-Phenoxybenzyl	(2) 3.2 (0.054)	0.17 (0.016)	D	D		
XX	Methoxy	Chloro, fluoro	3-Phenoxybenzyl	(2) 0.2 (0.007)	0.05 (0.0044)	D	B		
XXI	Ethoxy	Dichloro	N-tetrahydrophthalimidomethyl	(5) > 32 (> 32)	—	—	—		
XXII	Chloro	Dichloro	5-Benzyl-3-furymethyl	(3) 0.28 (0.0044)	0.28 (0.005)	A	D		
XXIII	Ethyl	Dichloro	3-Phenoxybenzyl	(2) 0.35 (0.016)	0.7 (0.057)	D	D		
IV**	Permethrin			0.05 (0.009)	0.011 (0.00043)	A	A		

*Mortality in µg per female insect. Details in caption to Fig. 1.

**Structure, Fig. 16.

LD₅₀ 0.52(0.01)

26.6(0.045) µg per insect

tetramethrin (V, Fig. 1b) formed an inactive ester with our acids (XXI, Table 1).

To further verify the extension of our DDT model to the combined structures we tested the 'apex' group dimethyl-oxetane³ which in the diaryl series yielded an active insecticide (XXIV).

Combined with the 3'-phenoxybenzylcarboxylate moiety (XXV), this derivative exhibited considerable insecticidal activity against the housefly, when potentiated with a synergist (value in parentheses). The high synergistic ratio for this compound is evidence of its ready degradation in the insect.

Physicochemical properties of the compounds were also investigated. This study was prompted by a report¹⁶ that in pyrethroid derivatives changes in lipid solubility (as determined by the partition coefficients) have a marked effect on their insecticidal activity. We determined the partition coefficients (log *P* octanol/water) for our compounds and their models, using a modified method of McCall¹⁷. With the exception of a limiting value of log *P* = 4.5 below which all derivatives lost their insecticidal activity, we found no significant correlation between the partition coefficients and the biological activities.

While mammalian toxicity of the new insecticides was not studied extensively, the dichlorocyclopropane VI (Fig. 1c) was tested for acute toxicity (mouse, subcutaneous injection in olive oil). At the practical dose-level limit of 16,000 mg per kg body weight, no symptoms of toxicity were detected in the 2-week observation period.

In conclusion we have demonstrated the possibility of structural combination between DDT isosteres and pyrethrum-based insecticides. This combination influences the

neurophysiological properties of the compounds and changes the sensory response in insects. This is shown by the increase in the insect behavioural property, for example, repellency from treated surfaces. The indicated low mammalian toxicity, ease of preparation¹⁸, and broad spectrum of activity (current insect spectrum expansion tests) should stimulate the development of a new range of safe insecticides. The demonstrated link between DDT and pyrethrum insecticides will enable the advancement of the study of the mode of action of both classes of compounds.

Patents of use and preparation of the new compounds are assigned to the Commonwealth Scientific and Industrial Research Organization. We thank the following for their help with this project: Mr K. Rihs, Dr T. Spurling, Mr A. Pompe, Dr B. Poppleton, Dr E. Shipp, Dr D. Smith, Mr G. Smith and Dr L. B. Barton Browne. Mammalian toxicities were determined at the Commonwealth Serum Laboratories, Melbourne.

G. HOLAN
D. F. O'KEEFE
C. VIRGONA
R. WALSER

Division of Applied Organic Chemistry, CSIRO,
P.O. Box 4331,
Melbourne, Victoria 3001, Australia

Received 19 October 1977; accepted 13 February 1978.

- Holan, G. *Nature* **221**, 1025-1029 (1969).
- Holan, G. *Bull. Wld. Hlth Org.* **44**, 355-362 (1971).
- Holan, G. in *Environmental Quality and Safety* (ed. Coulston, F. & Korte, F.) 359-364 (G. Thieme, Stuttgart, 1975).
- Elliott, M. *Bull. Wld. Hlth Org.* **44**, 315-324 (1971).
- Elliott, M., Farnham, A. W., Janes, N. F., Needham, P. H. & Pullman, D. A. *Pestic. Sci.* **6**, 536-542 (1975).
- Bercken, J. van den, Akkermans, L. M. A. & Zalm, J. M. van der *Eur. J. Pharmacol.* **21**, 95-106 (1973).
- Virgona, C., Holan, G. & Shipp, E. *Pestic. Sci.* **7**, 72-74 (1976).
- Narahashi, T. *Pestic. Sci.* **7**, 267-272 (1976).
- Burt, P. E. & Goodchild, R. F. *Pestic. Sci.* (in the press).
- Novak, J. & Farkas, J., W. German Patent, 1154463 (1963); *Coll. Czech. Chem. Commun.* **26**, 2090-2092 (1967).
- Shell Internat. Research Co., Netherlands Patent Appl., 7212973 (1973).
- Volkov, Yu. P. & Filonova, I. A., USSR Patent 210 143 (1968); *Khim. Farm. Zh.* **2**, 22-26 (1968).
- Sagar, W. C., Monroe, R. E. & Zabik, M. J. *J. agr. Food Chem.* **20**, 1176-1179 (1972).
- Barlow, F. *et al. Pestic. Sci.* **2**, 115-118 (1971).
- Staudinger, H. & Ruzicka, L. *Helv. Chim. Acta* **7**, 177 (1924).
- Briggs, G. G. *et al. Pestic. Sci.* **7**, 236-240 (1976).
- McCall, J. M. *J. med. Chem.* **18**, 549-552 (1975).
- Holan, G. & Walser, R. *W. German Patent Offen.* 26 43 189.5 (1977).
- Wood, J. C. *Proc. 7th Br. Insect. Fungicide Conference*, 273-279 (1973).
- Barton Browne, L. & Kerr, R. *Entomol. exp. appl.* **10**, 337-346 (1967).

reviews

Measurement of small forces

C. W. F. Everitt

Measurement of Weak Forces in Physics Experiments. By V. B. Braginsky and A. B. Manukin, edited by David H. Douglass. Pp. 153. (University of Chicago Press: Chicago and London, 1977.) \$10; £7.

THE science of measuring small forces began in the 1790s with Henry Cavendish's experiment to determine the gravitational constant and 'weigh the Earth'. Cavendish measured accelerations on lead spheres suspended from a six-foot long torsion arm to 10^{-10} g, an advance of five orders of magnitude over any previous work. His paper of 1798 still makes exciting reading. During the nineteenth century the torsion balance was applied in other gravitational experiments, notably those of R. von Eötvös and C. V. Boys. Boys reached acceleration levels of 10^{-13} g, and measured torques a factor of 10^7 lower than Cavendish's. The limits to measurement were also studied by electricians from Weber and Kelvin on, especially in the design of galvanometers. In 1925, W. J. H. Moll and H. C. Burger (*Phil. Mag.* 50 (sixth series), 624; 1925) noticed in the recorded output of a very stable galvanometer a small jittering motion, which they attributed to micro-seismic disturbance, but which G. Ising (*Phil. Mag.* 1 (seventh series), 827; 1926) identified as Brownian motion of the suspended body. The corresponding statistical noise limit in an electric circuit, the Johnson noise, was discovered by J. B. Johnson and interpreted by H. Nyquist one year later, in 1927. The great progress in measurement theory from then on owed much to Nyquist and to radar research in World War II. Torsion balances and gravity seemed forgotten, except by geophysicists, until R. H. Dicke, R. V. Pound, J. Weber and their respective associates began a revival in the late 1950s of Cavendish-style physics.

The long history makes it surprising that V. B. Braginsky and A. B. Manukin's *Measurement of Weak Forces in Physics Experiments* should be the first book on the subject. It is divided into three parts. The first analyses limits on the detection of small forces. The second gives the most complete account yet available of the equivalence principle experiment carried out by Braginsky and Panov in 1971, along with a brief discussion of drag-free satellite technology. The third discusses techniques for observing gravitational radiation, including both Weber bars and

other more recent suggestions.

In measuring small forces on a suspended body one can integrate below the Brownian oscillations. The limit is the force fluctuation, given by what is usually called the Nyquist formula, though it should be called the Einstein-Smoluchowski formula

$$\langle F^2 \rangle = 2M\beta kT/\tau$$

where M is the mass of the test body, β the damping coefficient, kT the thermal noise and τ the time of observation. For a mechanical oscillator β may be expressed as the reciprocal of the time τ_m^* of damping of oscillations, which is by definition $Q\tau_0/\pi$, where Q is the quality factor and τ_0 the natural period. Hence the common statement that measurement sensitivity is increased by increasing the Q of the system—a remark which underlies Braginsky's influential suggestion that gravitational antennae may be improved by using high Q sapphire crystals in place of Weber's aluminium bar. As the squared fluctuations are proportional to T/Q , there has been a nice debate about whether one gains more by working toward lower operating temperature or higher Q and about whether a small bar of high Q will be better than a large one of lower Q . The case for high Q is argued in chapter 9 of this book, which includes theoretical formulae and experimental data on the Q values of sapphire crystals as a function of temperature. A typical aluminium bar has a Q of 10^5 . The data reproduced by Braginsky and Manukin show values for Q of 10^9 for a 10-kg sapphire crystal at 2K; the theoretical limiting Q at that temperature is 10^{17} .

To measure the motions of a test mass, whether a torsion balance or a Weber bar, require a position sensor. This sensor will have intrinsic noise, which adds to the total noise output and also reacts back on the test mass. Procedures for analysing such effects by considering current and voltage noise in the equivalent circuit have long been known to electrical engineers, but the back reaction (due to current noise) was negligible in the room temperature antennae of Weber and received little attention from physicists working on gravity experiments before Braginsky. Braginsky's approach is sufficiently unorthodox from the electrical engineer's viewpoint for the connections to be obscure. Whether it is the worse for that may be a matter of taste.

sion begins (chapter 3) with the change in period of a mechanical oscillator from read-out reaction due, for example, to the force between the condenser plates in a capacitive read-out or the radiation pressure in an optical lever. Experimental results are given for an optical system with mirrors at the end of a 12-cm torsion arm. The authors then point out that any read-out has damping, which will increase the total dissipation entering equation (1). In chapter 4 they discuss quantum properties of a macroscopic oscillator and in chapter 5 the 'optimum strategy for measurement of small displacements'.

The chief interest of these calculations is in the design of gravitational antennae. For a torsion balance with optical read-out, the additional noise from random impacts of photons only become significant at temperatures below $6.9h/k\tau_0\epsilon^4$ where τ_0 is the natural period and ϵ the 'optical efficiency' of the system—which means for $\epsilon \approx 1\%$ and $\tau_0 \approx 10$ s a temperature below 0.1µK (C. W. F. Everitt, in *Proceedings of First Marcel Grossman Meeting on General Relativity*, ed. R. Ruffini, North-Holland: Amsterdam, 1977). At low temperatures an optical lever faces the objection that the light beam heats up the mirror, and this heat can only be removed by exchange gas. The additional gas damping is such that at temperatures below 80K, the limiting torque is higher than at room temperature unless the period of the balance exceeds about 12 h. Surprisingly the limiting torque, even at optimised pressures, scales only at $T^{1/2}$. Cryogenic operation brings only a meagre gain in noise performance to the torsion balance.

For gravitational antennae the quantity to be detected is not power or torque but a pulse of energy. The impulse sets the bar ringing; the motions are then coupled through a transducer to a low noise amplifier A. The best linear transducer is a parametric amplifier which takes in power from the bar at its natural frequency ν_B and multiplies it up to the operating frequency ν_A of A. The maximum power gain then is, according to the Manley-Rowe relationships ν_A/ν_B . The sampling time τ_s for measuring the signal should be short (though longer than the duration of the gravitational pulse) to minimise the disturbance of the ringing bar from subsequent thermal fluctuations. For a bar connected by a parametric transducer of gain G_p to a conventional linear amplifier

of noise temperature T_A , R. P. Giffard (*Phys. Rev.* **D14**, 2478; 1976) has calculated a minimum detectable pulse energy

$$\delta E < 2k\epsilon \left[\frac{\pi T_B \nu_B}{Q_B} + \tau_c + \frac{T_A}{G_p} \right]$$

where ϵ is the reliability factor defined on p5 of Braginsky and Manukin's text. An efficient antenna must have both terms in the bracket small. From the Manley-Rowe relationship, T_A/G_p is greater than $T_A \nu_B/\nu_A$. The best equivalent temperature with existing amplifiers is 100 μ K, which means for an antenna at 4K there is at present no advantage in having Q_B higher than 10^6 .

The argument from this point turns on two questions. First, how far can the performance of linear amplifiers be pushed? To this the answer is to the quantum limit, or equivalent temperature of 0.1 μ K, three orders of magnitude below the present state of the art. The second question is whether any device can be made to beat the quantum limit.

It is characteristic of both the strength and weakness of Braginsky that he leaps over—or rather tunnels through—the mundane barrier exposed by Giffard's formula and rushes on to the more spectacular work of beating the quantum limit. No-one reading pp44–48 of the text is likely to imagine that the signal might have to do anything so dull as pass through a noisy amplifier. The quantum limit is derived in chapter 4. Braginsky's "quantum non-demolition" scheme for measuring the energy state of an electromagnetic oscillator is described on pp48–53. Since the publication of this text, W. G. Unruh (*Phys. Rev. D*; in press) has shown that Braginsky's original scheme, being linear in the field, does not work. An alternative suggested by Unruh, having response quadratic in the field, should work in principle, though the difficulties are formidable.

The shortness of Braginsky and Manukin's text makes it unfair to expect detail. One section that does give new details is the account (chapter 6) of Braginsky and Panov's 1971 experiment on the equivalence of gravitational and inertial mass. This gave a quoted error of 1 part in 10^{12} as compared with 3 parts in 10^{11} for the experiment of Roll, Krotkov and Dicke. The result has been accepted by theorists but some experimentalists, including myself, have expressed doubts, particularly concerning the stability and resolution of the optical read-out (Everitt, *op. cit.*). The new details are interesting but do not set to rest all doubt. Of course, even if the sceptics were right and the limiting resolution no better than Dicke's, it is an achievement to have gone so far with such simple apparatus.

Every experiment on small forces from Cavendish's on, has required an attack on two problems: disturbances on the test

body and measurement errors. Braginsky has contributed so many ideas about measurement that it is odd to find in one of his proposals no discussion of the decisive measurement problem. Brevity is the cause, but it is a misleading brevity. In chapter 7 (p82) the authors suggest applying a drag-free satellite in Earth orbit to measure the time rate of change of the gravitational constant. On Dirac's well-known hypothesis, $\dot{G}/G$ may be expected to be of order $3 \times 10^{-11} \text{ yr}^{-1}$. The idea is to look for a change in orbit period over the year. If the compensation of non-gravitational disturbances is $10^{-11} g$ (the level reached in the DISCOS controller of the TRIAD satellite) their effect will be smaller than that from the predicted change in G . But what of the measurement problem? To determine the period to 3 parts in 10^{12} (10% of the predicted effect) one would have to measure the in-track position to 0.1 mm per orbit or 5 m yr^{-1} . Doppler tracking gives some-

thing like 15 m per orbit and there are geophysical perturbations from the Earth's higher harmonics of several hundred metres which cause model-dependent variations in the expected position from orbit to orbit having uncertainty of order 50 m. No experiment with a drag-free satellite can sort out the time variation in G from these.

Such criticisms should not deter anyone from reading this pioneering work or coming to grips with the acute minds of Braginsky and his co-author. Indeed, despite the obscurities of style I found some of the approaches more accessible than a conventional treatment might have been. This is a book everyone interested in gravity experiments and the measurement of small forces will want to buy. It is pleasantly inexpensive. □

C. W. F. Everitt is Adjunct Professor to the W. W. Hansen Laboratories of Physics, Stanford University, Stanford, California.

National socialism and German physics

Scientists under Hitler: Politics and the Physics Community in the Third Reich. By Alan D. Beyerchen. Pp. 304. (Yale University Press; New Haven, Connecticut and London, 1977.) \$17.50; £13.35.

THE present book deals with a most interesting period in the history of modern science, with the impact of National Socialism on the German physicists. Although this may seem to be a somewhat circumscribed subject, its importance lies in the effect which it had on the inability of the Germans to develop a nuclear weapon. Moreover, the author's thorough assessment of the direction of science in Germany, although frequently discussed before, deserves note. The present book can be regarded, within limits to be discussed presently, as a thorough and careful study of the problem as a whole.

There is an opening chapter dealing with pre-Hitler science in Germany which is followed by accounts of the dismissal of the scientists. Chapters 6 to 8 are devoted to "Aryan Physics" and these are probably the most important ones, not because the subject matter itself is of any significance but because it has been largely neglected, due to its patent idiocy, by other historians of science. Although it has been regarded, as it should be, with nothing but ridicule by respectable scientists, it offered a happy hunting ground for the careerist nonentities of the Third Reich. The author has dealt in detail and with ample documentation with this scientifically unrewarding phase of physics

in Germany. Even if all the relevant events in the lives of Phillip Lenard and Johannes Stark had been recorded before, nobody has considered them worth more than a few passing remarks. By descending into the morass of their small-minded intrigues with equally small-minded and insignificant party bosses, however, the author has recreated for us a world in which these childish antics could determine the fate of large research institutions and even the conduct of the German scientific war effort.

To the German physicist, the subservience of the whole scientific apparatus and the serious regard with which the most harebrained utterances, provided they seemed to have the approval of some party boss, were treated, nowadays appear like a comic strip but they were very serious at the time. It now seems incredible to use that these utterances by completely insignificant and untutored nonentities should not only have been heeded but have been regarded with all outward signs of a superior intelligence. And this on the part of outstanding scientists who were the undisputed experts in the field as everyone who was only roughly aware of the true circumstances must have known. What shakes the reader is not only the quite incompetent idiocy of the arguments put forward in discussion on these, but the earnest effort of the true scientists to smuggle secretly into their observations a much watered-down mention of the relativity principle, never to be mentioned by its true name. They must have hoped against hope that some day soon reason must return and something might be saved out of a disaster—which never came.

Looking at it after more than thirty years will seem even to a scientist who

lived through these times like a nightmare that only deserves to be forgotten as soon as possible. The smooth and learned narrative of Alan D. Beyerchen, however, lifts those years of terror and insanity out of a time into which he was not yet born, into an uneasy present which can treat the Weimar years as an admittedly unpleasant but respectable phase of democracy. Is that really true? Or does the real truth lie much nearer, in a Prussia that for a few years after World War I had played at sham democracy but then had to give up the charade enough to embrace Nazism and play again at real war? It is a question that still awaits its answer, of which the best we can hope for will be that it now has become less important.

The cataclysm out of which Germany seems to have miraculously arisen into the semi-sanity of the fourth Reich may deceive a certain number of post-War historians in order to provide a hurriedly reconstituted picture of a Germany returned to sanity. It remains an open question and one which this reviewer would be likely to give a different answer than the smooth and reassuring one provided by Dr Beyerchen. To open this question, however, cannot be our task and we must deal with the matter in hand: the momentous happenings within the third rank in the Third Reich.

The first and most serious confrontation took place in 1938 and concerned Heisenberg. Three attacking articles appeared in *Das Schwarze Korps* under anonymous authorship and were clearly aimed at Heisenberg. They were definitely dangerous and Heisenberg's defence—or rather his counterattack—was typical. It consisted in making use of an old family acquaintance between Heisenberg's grandfather and the SS-leader's father, dating from the time when both of these had been secondary school principals in Munich. Old Frau Himmler added—somewhat uncertainly—that some of Himmler's friends were, to her mind, slightly unpleasant but she promised to look into it herself. This she evidently did. Helped by the famous aerodynamicist Prandtl, Himmler wrote on the same day of his talk with Prandtl to Heisenberg, saying that he had the latter's case investigated and decided that the *Schwarze Korps* was to be forbidden to mount further attacks on him. To this, however, a caution was appended: to make in future clear to his readers that a definite distinction had to be made between recognition of results and the personal attitude of the researcher.

Heisenberg had clearly been playing with fire but his tactics of counter-attacking had evidently been successful. Stark had been at first the driving force behind the *Schwarze Korps*

attacks but he was, of course, not a match for Himmler himself and, but for a few futile attacks that petered out, it was the end of Stark (Lenard having given up in old age, even earlier). His eightieth birthday coincided with the decisive Seefeld conference in November, 1942. At this meeting, it had become clear to the German physicists that, in spite of staggering successes in the field, the War might be lost. The time for playing at war was over and those who considered the nature of the final end did not like it, although nobody was sure what it would be.

Only the strategists of the Third Reich kept on playing blind man's buff. Sticking faithfully to his brief, Beyerchen passes over much that would be interesting and holds on to the ridiculous and seamy side of Hitler's War, with its small intriguers and insignificant party politicians who blew themselves up like bullfrogs. It is perhaps the most fitting epitaph for a world that had gone crazy beyond redemption. Nobody had liked, although everybody knew about, the trains full of human beings that rolled east. An ever increasing number of Germans looked away. In all this human misery mounting month by month coming nearer and nearer, the German scientists were

pinning their last hopes on a V_1 , V_2 and perhaps an atomic V_3 . Perhaps, some of them hoped that in 30 years their enemies would find the whole thing slightly entertaining, if not downright comic. The world has a fortunate habit of using present uncertainties and trials as a diversion from the horrors of the recent past.

The author has, in addition to a compendious index, added lists of the relevant literature, taking up an appreciable volume of the whole book. As these contain a wealth of relevant information, it is only fair that their existence should be mentioned in any review of his book. There is, first of all a list of 839 references some of them disclosing a fair amount of detail. This is followed by a list of 10 papers and 11 dissertations and manuscripts as well as one of 11 tape-recorded interviews. The appendix is completed by lists of 107 books and pamphlets as well as of 84 articles having a direct bearing on the problem discussed in the book proper. It is hoped that this will give the reader some idea of the effort lavished on a passed and rapidly forgotten phase of our scientific heritage.

Kurt Mendelssohn

Kurt Mendelssohn is Emeritus Reader in Physics at the University of Oxford, UK.

Big bangs and black holes

Space, Time and Gravity: The Theory of the Big Bang and Black Holes. By Robert M. Wald. Pp. 131. (University of Chicago Press: Chicago and London, 1977.) \$11; £7.70.

WITH big bangs and black holes now firmly established in the public consciousness, the demand for concise, elementary and above all competent introductory books has reached new heights. The route to these 'fun' topics is through the theories of relativity and gravity, and many authors have searched for ways of softening the blows of these most abstruse and mathematical of subjects by invoking sweeping analogies, super-simplified discussions, elaborate diagrams or sheer intellectual sleight-of-hand. The truth is that a real understanding of the physical properties of black holes or the various cosmological models can only be achieved by several years of study of advanced mathematics. Semi-popular books can only have an appetite-whetting function and may seriously mislead the reader by painting too superficial a picture.

Having said this, it is encouraging that a young American scientist of international reputation, who himself has contributed much to the subject, has

produced a very readable and accurate account of modern relativity physics for the layman within the unavoidable constraint of almost no mathematics. Based on a course of lectures he gave at the University of Chicago in 1976, Dr Wald presents an introduction to the basic concepts necessary for an understanding of black holes and cosmology in a language that most laymen will be able to follow, if not to fully appreciate.

The layout of the book is standard: in the first three chapters the groundwork of relativity is laid, with the usual topics such as space-time diagrams, non-Euclidean geometry, discussions of clocks, bending, and so on, disposed of methodically and briefly but competently. The pay-off for this investment starts at chapter 4 when a discussion of cosmology follows. The language used here is more geometrical than other books at this level, which reflects the geometrical rather than physical approach adopted in describing the theory of relativity.

The author first explains the importance of the fact that the Universe on a large scale is roughly uniform and in a state of universal expansion. He then goes on to describe cosmic dynamics and the big bang itself. At this stage readers will naturally focus on the question of whether the big bang really represents the creation of the Universe in some sense. Dr Wald explains that general relativity admits so-called space-time singularities

and mentions the powerful theorems proved recently by Stephen Hawking and Roger Penrose, which indicates that they occur in a wide range of conditions. If the Universe was exactly uniform in density, then such a singularity at the beginning of the big bang would indeed represent a boundary of the physical world and a place of infinite compression. The author points out, however, that the theorems do not require this situation when departures from exact symmetry occur, as they do in the real Universe. Nor do we expect general relativity to apply unchanged all the way down to the singularity, because quantum gravity effects will presumably dominate at sufficiently early moments. Thus, he leaves open the question of whether the Universe existed before the big bang.

The remainder of the book deals with black holes. First, a description is given of how a star is constituted and burns, and the mechanism whereby it eventually fails to support itself internally against the crushing weight of its own material.

After a brief account of supernovae and neutron stars, the reader is introduced to black holes, with all their exotic and enigmatic properties. Dr Wald pays more attention than most authors to the intriguing question of cosmic censorship—can gravitational collapse cause the occurrence of space-time singularities that are not hidden inside black holes, that is, are naked? Drawing on some theoretical studies, he concludes that in all probability a 'cosmic censor' is at work to prevent us from viewing naked singularities.

A whole chapter is devoted to energy extraction from black holes that rotate, and a discussion of the efficiency limits on these processes. A further chapter deals with the astrophysics of black holes, for in spite of the wide attention they have received there is no clear observational evidence for them. Dr Wald explains how difficult it is to actually spot one, but draws the personal conclusion that the X-ray source known as Cygnus X-1 does contain a black hole.

The final chapter brings the subject

right up to date with an account of the quantum process of black hole evaporation discovered by Stephen Hawking in 1974. No attempt is made to describe the mechanism of the process in detail, but some of its consequences are explored, particularly its significance in providing a basis in thermodynamics for black hole processes.

In summary, I found this a well written, entertaining and authoritative book. It is extremely concise, so the reader should not expect anything other than a brief and occasionally patchy introduction. I see the book as excellent spare time reading for non-specialist scientists, good stimulation for school children with a scientific curiosity, and a helpful introduction for laymen and students to a subject of continuing scientific interest.

Paul Davies

Paul Davies is Lecturer in the Department of Mathematics at King's College, University of London, UK.

Digital image processing

Digital Image Processing. By Rafael C. Gonzalez and Paul Wintz. Pp. 431. (Addison-Wesley: Reading, Massachusetts, 1977.) Hardback \$29.50; paperback \$19.50.

In collaboration with different co-authors, Professor Gonzalez is writing a series of three introductory textbooks on pattern recognition and image processing. The volume under review is the second of these and is a direct rival of two volumes that have recently appeared with similar titles (reviewed in *Nature* 264, 142, 1976; and 268, 275, 1977). The contents of Gonzalez and Wintz's book are very similar to those of Rosenfeld and Kak's volume, but whereas Rosenfeld has made major contributions to both image processing and pattern recognition, the present authors have published comparatively little on most of the topics covered in this volume. Although they discuss all the material one would expect, their book is less useful than the rival texts in that it is difficult to grasp the role of any particular procedure, to place it in the context of contemporary attempts to process images in sophisticated ways. It will be even more difficult for the student to absorb some of the material, for the authors seem deliberately to have avoided analogies with other subjects that could have been helpful. The Hotelling transform (the discrete analogue of that of Karhunen and Loève) is a case in point: we are told, as though it were surprising, that pre- and post-multiplying a matrix by the matrix of its eigenvectors produces a

diagonal matrix. Yet a large part of image processing theory is applied matrix algebra, the elements of which are indispensable for any would-be reader of this book (or indeed any image processing book). Moreover, the student is likely to have met similar reasoning in quantum mechanics courses. In short, it is hard to imagine that anyone unaware of such an elementary result would embark on this text.

After introductory chapters on fundamentals—system requirements, sampling and film properties, in particular—separate chapters are devoted to transforms, image enhancement, restoration, coding, and image segmentation. The best of these is the chapter on image encoding, in which the authors success-

fully convey the reasons why the many different types of coding are necessary.

Digital Image Processing is a well illustrated, rather pedestrian text, containing plenty of examples and much rather wordy explanation. It will probably be welcomed by students with an examination to pass, but graduate students planning to continue in this field would be better off with the book by Rosenfeld and Kak. Incidentally, the hardback is probably worth the extra \$10, as the paperback is very flimsily bound.

P. W. Hawkes

P. W. Hawkes is Maître de Recherches in the CNRS Laboratory of Electron Optics, Toulouse, France.

Water and climate

Climate and Water Supply. (US National Research Council: Washington, DC, 1977.) Paperback \$7.75.

THE US National Research Council is producing a series of "Studies in Geophysics" on topics such as geophysical prediction, the upper atmosphere, energy and climate, and the present volume on water and climate. The studies vary in size and scope, with in this case a rather sharp focus on the water supply problems of the contiguous states of North America, which must inevitably reduce the usefulness of the volume compared with, say, the recent publication on energy and climate.

Nevertheless, some attempt is made to set these water supply problems within the broad framework of the developing

understanding of climatic change, with contributions from Stephen Schneider and Richard Temkin on "Water Supply and the Future Climate" and from Charles Stockton on "Interpretation of Past Climate Variability from Paleo-environmental Indicators". For readers outside the US, the result is best seen as a conveniently packaged and neat example of the application of studies of the changing climate to the specific problems of water supply in one part of the world; as such, the volume could be particularly useful in a teaching context for courses in geography or the Earth sciences. The volume is also a restrained and timely antidote to some of the more hysterical writings on climatic change that are now proliferating.

John Gribbin

John Gribbin is a Research Fellow of the Science Policy Research Unit, University of Sussex, UK.

Practical applications of the shell model

Shell-Model Applications in Nuclear Spectroscopy. By P. J. Brussaard and P. W. M. Glaudemans. Pp.452 (North Holland: Amsterdam, New York and Oxford, 1977.) Dfl.160; \$65.50.

THE shell model is the most basic and comprehensive of all nuclear models, and in the past three decades it has been developed to a high degree of sophistication and extensively used to analyse nuclear spectroscopic data. Nucleons interact strongly, so it is at first sight rather surprising that they can be assigned independent orbits inside the nucleus, but this comes about because of the Pauli Exclusion Principle that forbids many collisions that would otherwise occur and also the rapid fall-off of the nucleon-nucleon interaction with distance. Early calculations did not give the observed sequence of magic numbers of nucleons corresponding to closed shells, but this was remedied with the addition of the spin-orbit potential in 1949. Thereafter, the shell model has proved astonishingly successful in accounting for a wide range of nuclear reaction and structure data.

The book by Brussaard and Glaudemans is a welcome addition to the already considerable literature on the shell model. It is a comprehensive textbook at the level of the first-year graduate student, with detailed chapters on the more important aspects of the model. The formalism is derived in detail, so that it is useful for reference as well as for self-study.

After an introductory chapter on the foundation of the shell model, the formalism for one- and two-particle systems is described, and perturbation theory is applied to calculate binding and excitation energies. Shell model calculations require the evaluation of matrix elements of interaction operators between many-particle states, taking account of configuration mixing and using angular momentum coupling and coefficients of fractional parentage. This is described in detail and applied to the cases of interacting particles in one and two orbits.

The dimension of the configuration space in shell model calculations increases very rapidly with the number of active particles. It is therefore usual to truncate the space and use an effective interaction the parameters of which are adjusted to fit selected data. A particularly successful interaction is the surface delta interaction, and this is applied to several nuclei.

Shell-model wavefunctions are used

to calculate the cross-sections of one- and two-nucleon transfer reactions, and this is one of the principal ways of determining spectroscopic factors. Electromagnetic transition operators and rates are also calculated with shell model wavefunctions, as well as electric and magnetic multipole moments and beta decay rates. All these subjects are described, with many practical applications.

Additional chapters summarise the second quantisation formalism, realistic and effective operators, the applications of Wick's theorem, the Lanczos method, Coulomb energies, spurious states and the particle-vibrator model. There are many appendices on coupling coefficients, the Wigner-Eckart theorem, reduced matrix elements, and

tables of 3-j and 6-j symbols, fractional parentage coefficients and the matrix elements for some simple potentials.

This book will certainly be very useful to all who have to learn about the shell model and make practical calculations. Some references to the original work are given, but the value of the book would have been increased by a more comprehensive list. The section on Coulomb energies is too brief and contains no mention of the Nolan-Schiffer anomaly. The book is excellently printed and the price is correspondingly high.

P. E. Hodgson

P. E. Hodgson is Lecturer in Nuclear Physics at the University of Oxford, UK.

Physics in perspective

A Perspective of Physics. Edited by Rudolph Peierls. Pp. 243. (Gordon and Breach: London, New York and Paris, 1977.) £20.

THE rate of progress of significant research in physics remains very high. There is indeed no sign as yet of it falling off, despite the financial strains in the western world during the past few years. Under these circumstances it is essential to have continually available reviews and comments at different levels so that the subject does not become too over-specialised. The so-called *Comments Journals* dealing with nuclear and particle physics, solid-state physics, astrophysics, atomic and molecular physics, and plasma physics and controlled fusion, fulfill a very useful role in this respect. This particular volume goes further in presenting a selection of comments from all these areas, so that from them the reader may obtain some idea of the growing points in the subject as a whole.

For a volume of this kind to be a success it is essential that the selection be made by some distinguished physicist who does see the subject as a whole. No-one better than Sir Rudolf Peierls could be imagined for this task. Furthermore the value of the book is greatly increased by the masterly introduction by Sir Rudolf which attempts with great success to put the whole into perspective in a scant 26 pages. It is remarkable how well-balanced is this Introduction, covering as it does such different kinds of research ranging from very fundamental studies in particle physics and in astrophysics through solid-state and atomic physics, and involving a rich variety of

phenomena depending on known basic laws, to a highly technical subject such as controlled fusion.

There is great excitement today from the new, unexpected discoveries of new particles and new quantum numbers which can hardly be said to have simplified the basic description of Nature. On the other hand very promising developments towards unifying the theories of the weak, electromagnetic and strong interactions have been made. The weakest of all the interactions, the gravitational, is the one which dominates the exotic aspects of astrophysics in which white dwarfs and neutron stars are commonplace and black holes are on the brink of observational confirmation. In these areas of 'big' science the demands on technology are very severe, so that the most advanced technologies today are to be found in these fundamental research activities. Nevertheless there is a great amount of highly applicable work being done in the smaller sciences, again depending very much on new techniques. After a long period of uncertainty it now looks very probable that power from controlled fusion is a practical possibility, though still a distant one, and high morale exists among those working in this area.

Apart from the Introduction the book includes eight selections from *Comments on Nuclear and Particle Physics*, five each from *Solid-State Physics*, *Astrophysics* and *Atomic and Molecular Physics* and four from *Plasma Physics and Controlled Fusion*. The selection is a good one and should prove absorbing to a physicist anxious to know something of what is going on around him beyond his own direct involvement.

Harrie Massey

Sir Harrie Massey is Emeritus Professor of Physics in the University of London, UK.

Collective papers in cognitive psychology

Thinking: Readings in Cognitive Science. Edited by P. N. Johnson-Laird and P. C. Watson. Pp. 615. (Cambridge University Press: Cambridge, New York and London, 1977.) Hardback £17.50; paperback £5.95.

THE term *Cognitive Science* is used by the editors of this volume to refer to the new approach to "the scientific study of thinking" which has emerged from the cross-fertilisation, debates and occasional collaborations between cognitive psychology, linguistics and artificial intelligence (AI) during the past decade. The hallmark of the approach is its espousal of process models of human behaviour: postulated information-processing structures, typically cast in the form of computer programs, which suffice to generate the phenomena of interest. This approach contrasts sharply with the more traditional methodology in cognitive psychology which sees as its role the more functional task of establishing empirically what factors affect performance and in what ways. Questions about the underlying mechanism, if addressed at all, are discussed in terms deriving from neurophysiology rather than information processing.

The book contains 34 readings (three of them written specially) divided into seven sections: Problem Solving; Deduction; Conceptual Thinking; Hypotheses; Inference and Comprehension; Language, Culture and Thinking; and Imagery and Internal Representation. Each section begins with an extensive introduction to the topic and there is a ten-page Introduction to the book as a whole. Despite its interdisciplinary stance, in practice the book draws mainly from sources in cognitive psychology with occasional contributions from AI, philosophy and anthropology. Some 23 of the papers are by authors whose primary allegiance is to psychology and 6 are by people in AI (including one example of genuine collaboration across disciplines), leaving 5 from other fields. To the extent that it is of value to speak of 'themes' running through such a collection of papers, the most striking have to do with the pervasive role of inference in language comprehension and the difficulty people have in trying to draw conclusions from negative information. A secondary refrain concerns the fragile basis of Man's claim to rationality.

Whatever one feels about some of the individual choices, there can be no doubt that the collection as a whole is a stimulating one. Some of these papers have become minor classics, either because of their seminal quality (Shepard and

Metzler's original paper on mental rotation of objects, Huttenlocher's and H. H. Clark's contrasting analyses of three-term series problems) or because of their concise summary of an area of investigation (for example, Tversky and Kahneman's catalogue of human fallibilities in handling probabilistic information). The Genevan viewpoint is represented by extracts from a paper by Piaget dealing with intellectual development from adolescence to adulthood, and by a delightful piece by Karmiloff-Smith and Inhelder in which the careful, insightful observation and description of children's attempts to balance wooden blocks provides a salutary example to workers in AI and in mainstream cognitive psychology alike.

This collection was assembled specifically to meet the needs of students taking a new Open University course on cognitive psychology. It will be of relevance to all those who are interested in the impact that advances in linguistics and AI have had on psychologists' view of cognition. However, the book is not

without its defects. Apart from typographical errors, a couple of the papers have been 'edited' to the point of near-incomprehensibility. The distinctions between some of the seven sections are unclear. And the reader should be forewarned that the particular view of cognitive science which shapes the book is very much the editors' own. The emphasis is heavily biased towards the linguistic aspects of cognition, with other areas being painfully under-represented. In several places the editors come close to identifying 'thinking' with (verbal) 'reasoning', and when reading the introductions to each section this reviewer had several times to re-assure himself that he had not accidentally picked up a textbook on psycholinguistics. These *Readings* need to be supplemented by other readings to provide a balanced picture of modern thinking about thinking.

Richard M. Young

Richard Young is an SSRC Fellow in the Department of Artificial Intelligence at the University of Edinburgh, Scotland.

Superradiance rules the day

Cooperative Effects in Matter and Radiation. Edited by C. M. Bowden, D. W. Howgate and H. R. Robl. Pp. 388. (Plenum: New York, 1978.) \$45.

R. H. DICKE's famous (and obscure) paper on cooperative spontaneous emission of radiation, or "superradiance", was published twenty-four years ago. The present volume, representing the proceedings of a co-operative effects meeting at Redstone Arsenal, Alabama, in late 1976 consists for the most part of descriptions of various approaches to superradiant theory and experiments. The word superradiant has often been misused (even incorrectly applied to high-gain mirrorless lasing) but is now taken to mean either the coherent radiation generated by a system of atoms possessing an externally created macroscopic polarisation rather akin to a phased array of dipoles, or alternatively the incoherent fluctuating radiation produced by a system of excited atoms possessing no initial polarisation but instead developing a cooperative decay behaviour through quantum correlations. The word "superfluorescence" has been proposed by Rodolfo Bonifacio for this latter process which can be distinguished by its coherence properties and its directionality. The explosion of theoretical interest in the late 1960s and early 1970s in superradiance has been followed very recently by a number of experiments on far-infrared superradiance, described

here by Gibbs, Vrehen and others.

Leading problems in superradiance are whether or not single-pulse emission or "ringing" occurs, and how the initial coherence is established. The two leading theoretical camps are well represented here. The mean-field approach advocated by Bonifacio, with single-pulse emission and with no role being played by propagation, is described by Bonifacio and by Bullough. On the other hand, the computer-solution approach to the totally semiclassical coupled Maxwell-Bloch equations is described by Mike Feld and emphasises propagation and associated ringing of the pulsed emission. The discussion of the conference papers is published verbatim at the end of this volume and illustrates the heat generated by this topic. The interesting and difficult problem of the initial evolution of the system and the establishment of coherence is indicated as the outstanding difficulty left to solve.

Other papers in this volume are concerned with the free-electron laser experiment at Stanford and the theory behind it, the Dicke phase transition, and self-focussing in self-induced transparency. But superradiance obviously ruled the day. This is a stimulating volume, unusual in so clearly depicting the problems remaining to be solved and the conflicts between rival explanations. It is specialised and expensive, but certainly deserves to be in institutional libraries.

Peter Knight

Peter Knight is Jubilee Research Fellow in the Department of Physics at Royal Holloway College, University of London, UK.